COMPARATIVE ANALYSIS OF SOYBEAN (*GLYCINE MAX*) ACCESSIONS USING INTER SIMPLE SEQUENCE REPEAT (ISSR) AND RANDOM AMPLIFIED POLYMORPHIC DNA (RAPD) MARKERS

by

Sarah Alamri

Master of Science (MSc) in Biology

A thesis submitted in partial fulfillment
Of the requirement for the degree of
Masters of Science (MSc) in Biology

The School of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
School of Graduate Studies/École des études supérieures

Title of Thesis Titre de la thèse	COMPARATIVE ANALYSIS OF SOYBEAN (GLYCINE MAX) ACCESSIONS USING INTER SIMPLE SEQUENCE REPEAT (ISSR) AND RANDOM AMPLIFIED POLYMORPHIC DNA (RAPD) MARKERS
Name of Candidate Nom du candidat	Alamri, Sarah
Degree Diplôme	Master of Science
Department/Program Département/Programme	Biology
Date of Defence Date de la soutenance	April 28, 2014

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Kabwe Nkongolo
(Supervisor/Directeur de thèse)

Dr. Abdel Omri
(Committee member/Membre du comité)

Dr. Victor Clulow
(Committee member/Membre du comité)

Dr. Benoit Bizimungu
(External Examiner/Examinateur externe)

Approved for the School of Graduate Studies
Approuvé pour l'École des études supérieures

Dr. David Lesbarrères
M. David Lesbarrères
Director, School of Graduate Studies
Directeur, École des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

I, **Sarah Alamri**, hereby grant to Laurentian University and/or its agents the non-exclusive license to archive and make accessible my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that this copy is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

Abstract

Soybean (*Glycine max)* is an important crop in the world in terms of total production and usage. It is also among the least diverse species. The main objectives of the present study were 1) to determine differences between ISSR and RAPD marker systems in detecting genetic variation in soybeans and 2) to identify and characterize accession- diagnostic molecular markers in *G. max* accessions. Genomic DNAs from 108 *G. max* accessions from 11 different gene pools were analyzed using several ISSR and RAPD primers. The levels of polymorphic loci detected with the two marker systems were in general moderate and similar.. Overall, 82% of genetic distance values were above 0.40 based on ISSR analysis. However, RAPD data revealed that the accessions from different countries are closely related with 64% genetic distance values below 0.40. The dendrograms constructed with ISSR data revealed that the South Korean accessions formed an out-group while the RAPD analysis showed that accessions from Sweden were separate from the other 10 gene pools. One variety-diagnostic marker generated with ISSR 5 primer was identified in the accession *Kao Chien Tao* from China. This marker was cloned, and sequenced. Although RAPD and ISSR marker systems detected similar levels of genetic variability, they target different regions of the soybean genome, resulting in different clustering of the 11 gene pools indicating different genetic relatedness among them. This finding demonstrates the usefulness of both marker systems in assessing diversity and relatedness among *Glycine max* gene pools.

Keywords

Soybean, *Glycine max,* Genetic diversity, ISSR, RAPD, SCAR, Molecular markers

Acknowledgments

A major research project like this is never the work of anyone alone. Thank God for the wisdom and perseverance that he has been bestowed me during this research project. My thanks to Dr. Kabwe Nkongolo, for making this research possible, with his support, guidance, advice throughout the two years of my stay in Canada; to my lab mates for their technical assistance; and my parents, brothers, sisters, and my husband for their unconditional support. The financial support from the Ministry of Higher Education of Saudi Arabia is greatly appreciated.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Literature Review

1.1 Importance of Soybean

Soybean (*Glycine max*) is the most important legume in the world in terms of total production. The leading grain producers of this crop include the United State (46%), Brazil (20%), Argentina (13.5%) and China (9%); but significant production also occurs in India, Paraguay, Canada and Indonesia (Soya & Oilseed Bluebook, 2001; Chen Y., 2002). At nutritional level, soybean contains all eight essential amino acids necessary for human body to make protein (Fukushima, 2001). It also contains isoflavones and fiber,,that provide health benefits. It is the most recommended source of complete protein especially for vegetarians. Soy isoflavones and lecithin have been studied extensively for potential health benefits (Ye et al., 2012). Several studies have shown that isoflavones such as genistein have estrogen-like effects in the body, and for these reasons are called "phytoestrogens" (Ye et al., Wang Z.L 2012).

Soybean (*G. max*) has long been recognized as a high quality food because of its relatively high level of protein compared to other plants. Recent studies revealed that, even though soy proteins are of plant sources and typically lower in certain amino acids compared to animal proteins such as those found in eggs or cow's milk; it receives similar protein quality rating as egg or cow's milk protein (Anderson and Bush 2011). The discovery of very small and unique proteins in soy referred to as peptides increase its nutritional values. Examples of these unique peptides in soybeans include defensins, glycinins, conglycinins and lunasins, and all are now known to provide health benefits, such as regulation of blood pressure, control of blood sugar levels, and improvement of immune function (Bush et al., 2011).

1.2 Soybean Germplasm Collection

The cultivated soybean, *Glycine max* (L.) Merr., and wild soybean *G. soja* Sieb., and *G. soja* Zucc., both belong to genus *Glycine* and subgenus Soja. Both cultivated and wildtype soybeans are cross fertile. Maintaining and evaluating annual *Glycine* soybean germplasm and identifying and exploiting new genes in this germplasm are potentially beneficial to the success of soybean cultivar development (Chen Y., 2002).

The geographical distribution of wild soybean (*G. soja)* is limited to East Asia including China, Korea, Japan and the Far East of Russia. It is believed that the cultivated soybean (*G. max)* originated in China, was introduced to Korea, Japan and Russia and eventually spread around world. The International Plant Genetic Resources Institute reports that there are 129 institutions worldwide holding 149,613 accessions of *G. max* and 24 institutions with 9,538 accessions of *G. soja* (update 1999). Ten of 129 *G. max* collections maintain half of the accessions, and two of *G. soja* collections hold over 70% of the accessions. The most significant collections are from: Institute of Crop Germplasm Resources, CAAS (China); USDA Soybean Germplasm Collection; Asian Vegetable Research and Development Center (AVRDC); Institute of Agroecology and Biotechnology (Ukraine); N.I. Vavilov Research Institute of Plant Industry (Russia); Department of Genetic Resources National Institute of Agrobiology Resources (Japan); Crop Experiment Station Upland Crops Research Division (Korea) and Australian Tropical Crops Genetic Research Center (Chen Y., Nelson R.L. 1999).

China is the origin of soybean and has the largest collection of *G. max* and *G. soja* in the world. National soybean germplasm collecting efforts occurred in 1956, 1979 and 1990 (Chang et al. 1999; Chen Y., 2002). A total of 23,578 Chinese soybean accessions have been

documented in four volumes of the Chinese Soybean Germplasm Catalog (Wang, 1982; Chang et al., 1991, 1996, 1999; Chen Y., 2002). An additional 1,946 introduced lines, mainly developed cultivars and genetic stocks from the United States complete the total of 25,524 accessions stored in the National Plant Gene Bank of China. Collection catalogs were published in China in 1990 and 1994 documenting the largest *G. soja* collection in the world (6,172 accessions from China) (Li, 1990, 1994; Chen Y., 2002). These publications contain 14 items of information on the origin, descriptive characteristics, agronomic performance and seed composition (protein and oil).

The USDA Soybean Germplasm Collection at University of Illinois has the second largest soybean germplasm collection which has over 18,000 accessions of *G. max* and 1104 accessions of *G. soja* (Randall, 2001; Chen Y., 2002). Juvik et al., (1989) published the first USDA *G. soja* evaluation bulletin and described 200 accessions for maturity group, seed characteristics and plant traits. Fewer than 20% of *G. soja* accessions now in this collection have been evaluated. The Asian Vegetable Research and Development Center (AVRDC) in Taiwan has the third largest soybean germplasm collection with 12,505 accessions of *G. max* and 339 accessions of *G. soja* according to the International Plant Genetic Resources Institute report (Chen Y., Nelson R.L. 1999). Identification and utilization of diverse germplasm is critical to the success of any plant-breeding program. The knowledge of genetic diversity within the *ex situ* collections and the genetic relationships within the subgenus Soja are useful for parental selection and germplasm management (Chen Y., 2002).

1.3 Morphological Characterization

Phenotypic evaluation of soybean germplasm is a fundamentally important procedure for the management of crop germplasm collections and for determining genetic diversity. There are three types of data (descriptive, agronomic and seed composition data) documented within

the Chinese and USDA soybean germplasm collection (Wang, 1982; Chang et al., 1991, 1996; Nelson et al.,1987, 1988; Chen Y., 2002). Traits, described by colors, shapes, appearances or forms are involved in the descriptive data. Agronomic data consist of scored or measured traits such as lodging, shattering, mottling, seed weight, plant height and maturity data. Oil from the seeds contain have five major fatty acids: palmitic, stearic, oleic, linoleic and linoleic acid. The classification of these traits has been well established for evaluating *G. max* germplasm. These same traits are often used to evaluate *G. soja* germplasm, but because of the differences between the two species some different evaluation procedures are needed (Chen Y., 2002).

1.4 Genetic Variation

The genetic diversity of wild soybeans has been studied by various techniques including isozymes (Shimamoto et al., 1998; Abe 2000), RFLP (Abe et al., 1999; Abe 2000), SSR markers (Lee et al., 2008; Powell et al., 1996) and cytoplasmic DNA markers (Abe et al., 1999; Shimamoto et al., 1998; Xu et al., 2002). Based on haplotype analysis of chloroplast DNA cultivated soybean appears to have multiple origins from different wild soybean populations (Xu et al., 2002). In addition to SSR markers being used in genetic diversity studies they have also been used to study gene flow between cultivated and wild soybean in Japan and China (Jin et al. 2003; Kuroda et al. 2006, 2008). SSRs have been previously used to analyze and develop the main core collections for wild and cultivated mungbean (Sangiri et al., 2007).

Several studies on soybean have concluded that genetic diversity is lower compared to other species. This lack of genetic diversity is due in part to a small genetic base available for breeding, which has caused challenges for soybean breeders (Apuya et al., 1988).

1.5 Assessing Genetic Variation Using Molecular Markers

Many phenotypic characters of soybean are heavily influenced by the environmental conditions under which the plants grow. Protein or enzyme variation can be used to study genetic diversity of crop germplasm. However, the limited number of isozymes of proteins and enzymes can limit their usefulness. Polymorphic DNA markers can provide an ideal alternative method for evaluating genetic diversity in soybean germplasm (Apuya et al., 1988; Keim et al., 1989; Chen Y., 2002).

DNA markers can detect numerous genetic variants in crops based on the analysis of differential DNA sequences. Theoretically, they are nearly unlimited in numbers, presumably selectively neutral and can be organized into linkage maps (Thormann and Osborn, 1992; Chen Y., 2002). There are many reports of using molecular markers for evaluation of soybean germplasm. Soybean RFLP markers were first introduced in the late 1980s (Apuya et al., 1988; Keim et al., 1989; Chen Y., 2002). Apuya et al., (1988) explored the possibility of using RFLPs as a genetic marker and mapped 11 of 27 RFLP markers into four linkage groups using an F_2 population. Keim et al., (1989) applied 17 RFLP markers to 58 soybean accessions from the genus *G. soja* and found that the genetic diversity was least within cultivated soybean group but greater between *G. max* and *G. soja.* Amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) markers have been shown to be more polymorphic in soybean than RFLPs (Powell et al., 1996b; Chen Y., 2002) Maughan et al., (1996) evaluated 23 *G. max* and *G. soja* accessions with 759 AFLP fragments. They found that 36% were polymorphic across all genotypes. Within the group of *G. soja* accessions, 31% were polymorphic, but only 17% were polymorphic within *G. max* accessions.

RAPD markers have been shown to be a simple and effective means to evaluate variability in crop. Based on principal component analysis of RAPD data on 35 soybean lines, Thompson and Nelson (1998) established a core set of RAPD primers with high polymorphism in soybean. These 35 core RAPD primers have been used in other studies for genetic diversity analysis in soybean (Chen Y., and Nelson, 1999; Brown-Guedira et al., 2000; Chen Y., 2002). With RAPD markers, Thompson et al., (1998) evaluated 18 U.S. soybean ancestors and 17 selected accessions from the USDA Soybean Germplasm Collection. The clusters defined by the RAPD data corresponded to known pedigrees, origins and maturity groups.

The first demonstration of simple sequence repeat (SSR) allelic variation and heritability in a plant species was in soybean, and SSRs have been shown to be highly polymorphic in soybean (Akkaya et al., 1992; Diwan and Cregan, 1997; Chen Y., 2002). Akkaya et al., (1992) employed SSRs to evaluate the diversity of 43 ancestral and commercial cultivars representing the U.S. gene pool. They determined that, in general, SSRs with AT core motifs are most polymorphic in soybean, followed by those with ATT cores. As many as 26 alleles at a single locus have been found. Maughan et al., (1995) reported that 5 SSR markers detected a total of 79 alleles across 94 accessions of *G. max* and *G. soja.* With 43 more SSR alleles detected in *G. soja* than in *G. max*, the result again demonstrated the greater diversity in *G. soja.* Rongwen et al., (1995) detected 11 to 26 alleles at each of seven SSR loci in a diverse sample of soybean genotypes including U.S. cultivars, and introductions of *G. max* and *G. soja,* including Chinese landraces. Twenty-six alleles at a single locus detected in this study are the largest number of alleles at an SSR loci reported in soybean to date (Rongwen et al., 1995).

1.6 Description of ISSR and RAPD Markers

1.6.1 Inter Simple Sequence Repeat (ISSR) Marker

ISSR is a relatively new technology that is used to differentiate closely related individuals (Godwin et al., 1997). This marker system has been successfully applied to genetic analysis of plants. ISSR marker system accesses variation in the numerous micro-satellite regions dispersed throughout the genome (Semagn et al., 2006) and circumvents the challenge of characterizing individual loci that other molecular approaches require. ISSR involves amplification of regions between adjacent, inversely oriented microsatellites, using a simple sequence repeat (SSR) motif containing primers anchored at 3' or 5' end by two or four arbitrary, often degenerate nucleotides (Zietkiewicz et al., 1994). Microsatellites are very short (usually 10-20 bp) stretches of DNA that are hypervariable, expressed as different variants within populations and among different species. They are characterized by mono-, di- or tri-nucleotide repeats (AA, AG, CAG respectively) that have 4-10 units side by side. ISSR marker system is based on the use of 15-20 bp primers designed to be complementary to microsatellite sequences found throughout eukaryotic genomes. Therefore, this PCR based technique involves the amplification of DNA segments present between two identical microsatellites that are oriented in opposite directions (Reddy et al., 2002). ISSRs specifically target the di- and tri- nucleotide repeats a type of microsatellite that is characteristic of the nuclear genome (mono nucleotide are found in chloroplast genome (Queller et al., 1993).

The ISSR method provides an alternative choice to other systems for obtaining highly reproducible markers without any necessity for prior sequence information for various genetic analyses. ISSR method takes advantage of the ubiquitously distributed SSRs in the eukaryotic genomes. Because of those abundant and rapidly evolving SSR regions, ISSR amplification has the potential of revealing larger numbers of polymorphic fragments per

primer than any other marker system used such as RFLP or microsatellite. As the PCR reaction amplify the sequence between two SSRs, the PCR products generated reveal multilocus profiles which could be revealed on agarose or polyacrylamide gels.

The number of targeted genomic loci can be altered by designing primers of different specificities (Zietkiewicz et al., 1994). ISSR products can be easily excised from the gel and cloned or reamplified to be used as probes. This could be an alternative method of identifying microsatellites in genomic libraries (Zietkiewicz et al., 1994). ISSR analysis is faster and it amplifies and detects a greater number of bands per primer (Godwin et al., 1997). Study done by Godwin et al., (1997) found that ISSR markers reveal higher levels of polymorphism.

ISSRs provide great potential to determine intra-genomic and inter-genomic diversity compared to other arbitrary primers, since they reveal variation within unique regions of the genome at several loci simultaneously. Several properties of microsatellite such as high variability among taxa, ubiquitous occurrence, high copy number in eukaryotic genomes make ISSRs extremely useful markers (Morgante et al., 1996). They exhibit specificity of sequence tagged site markers, but need no sequence information for primer synthesis having the advantage over random markers. The multiple profiles generated by ISSR primers are highly polymorphic and as such are ideal for the study of genetic variability.

The basic premise of ISSRs is that primer annealing sites are distributed evenly throughout the genome such that the primer will anneal at two sites orientated on opposing DNA strands. If this is within an appropriate distance from one another, the region between two primers will be amplified through PCR. ISSRs are now being applied to natural populations to address issues such as hybridization. These studies have demonstrated the utility of this technique in a wide range of applications and plant families (*Asteraceae, Brassicaceae, Hippocastanaceae, Orchidaceae, Poaceae, Scrophulariaceae* and *Violaceae*).

Compared with other molecular marker such as AFLP and SSR, ISSR has its specific advantages including: 1) no prior sequence information required, 2) simple and quick operation, 3) amenable to laboratory level, 4) high stability, 5) abundance of genomic information, 6) use of radioactivity is not required, and 7) show high polymorphism. ISSR markers access variation in the numerous microsatellite regions dispersed throughout the various genomes (particularly the nuclear genome) and circumvent the challenge of characterizing individual loci that other molecular approaches require (Semagn et al., 2006; Sharma et al., 2008).

A number of advantages are associated with the ISSR multilocus technique. Firstly, ISSRs are universal in the sense that microsatellite repeats are found in every eukaryotic genome studied to data. This is most likely due to the longer lengths of primers which permit the use of higher annealing temperature which in turn, reduces non-specific binding and results in higher stringency (Bornet and Branchard, 2004; Qian et al., 2001). Finally, because each band corresponds to a specific DNA sequence by two inverted microsatellites, and the amplified products usually range from 100-2000bp long they can be detected by both agarose and/or polyacrylamide gel electrophoresis (Reddy et al., 2002).

1.6.2 Random Amplified Polymorphic DNA (RAPD) Marker

RAPD were first described by Williams et al., (1990). This was the first technique to amplify DNA fragments from any species without any prior sequences information. RAPD is a simple and easy method to determine genetic diversity and taxonomic identity of various species (Mishar et al., 2009). This technique has been widely applied in studies of genetic diversity and genetic structures of woody plant such as *Qurecus liaotungensis* (Liaiding oak), *Populus tacamoahaca* (Balsam poplar) Glycine (Soybean) and Tilia (Lime) (Congwen and Manzhu, 2006). The system involves the use of a single oligonucleotide of arbitrary sequence to prime

the amplification of template DNA by PCR. An oligonucleotide will prime amplification from a genomic template if the binding site on the opposite strand of the template exists within a distance, which can be traversed by the DNA polymerase (up to several thousand nucleotides).

The amplification with arbitrary primers is mainly driven by the interaction between primer, template annealing sites and enzymes (Semagn et al., 2006). Genomic polymorphisms at one or both priming sites result in the non-amplification of a band. RAPD are thus dominant markers and appearance of a band implies homology with the primer used. All other alleles at the priming site will be represented by absence of the band. Dominant RAPD markers resulting from insertions or deletions between priming sites and observed as different sized fragments amplified from the same locus, are detected rarely (Williams et al., 1990). A primer usually amplifies several bands, each originating from a different genomic location. The nature of the fragments amplified is influenced dramatically by the sequence of both primers and template.

RAPD usually uses a 10 bp arbitrary primer. Although the sequences are arbitrarily chosen, two basic criteria must be met: at least 50% Guanine-Cytosine content and the absence of palindromic sequences (Semagn et al., 2006). Primers as short as 5 nucleotides give more complex banding patterns requiring more sophisticated electrophoretic and staining procedures (acrylamide gels and silver staining). RAPD analysis results in the amplification of one locus and two kinds of polymorphism: the band may be present or absent, and the brightness of the band may be different. Band intensity differences could be due to low copy number or relative sequence abundance (Devos and Gale, 1992) and may serve to distinguish homozygote dominant individuals from heterozygotes, as less bright bands are expected for the latter. Ellsworth et al., (1993) indicated that the fact that fainter bands are generally

robust, varying degrees of primer mismatch may account for band intensity differences. As the source of the band intensity difference is uncertain most studies disregard scoring differences in band intensity (Semagn et al., 2006).

RAPD system does not require Southern blotting, development of species-specific probes or radioactive labelling. RAPD analyses can be conducted much more quickly and with fewer laboratory restriction and quantities of DNA required for analysis are 100 times lower (5-20 ng) compared to other markers. The dominant nature of the marker results in an inability to distinguish homozygotes and heterozygotes which is a limitation for some applications (Sharma et al., 2008). RAPD markers can be used for mapping in areas of the genome not accessible to other analyses due to the presence of repetitive DNA sequences. As with other genetic markers, some polymorphisms are easy to score, while others are ambiguous and not useful as markers. Ambiguous polymorphisms may result from poor discrimination by a primer between alternatives priming sites of slightly different nucleotide sequences.

The objectives of the present study were 1) to determine differences between ISSR and RAPD marker systems in detecting genetic variation in soybeans and 2) to identify and characterize accession-diagnostic molecular markers in *G. max* accessions from 11 different countries.

Chapter 2: Genetic Variation in Soybean (*Glycine max)*

2.1 Introduction

Diversities in many crops such as soybean have been documented based on morphological and agronomical traits (Lee and Kaltsikes, 1973; Ford and Ball 1991; Lu et al., 2002). Such characters are strongly influenced by environmental factors and the developmental stage of the plant. Since the early nineties, molecular markers such as SSRs, ISSRs and RAPDs have been developed as alternative methods and they are extremely effective in population genetics studies (Prasad et al., 2000; Ma et al., 2004). ISSR and RAPD primers were used in the present study to determine the genetic diversity and distance among *G. max* accessions collected from various regions around the world.

2.2 Materials and Methods

2.2.1 Genetic Materials

G. max varieties from 11 countries provided by The Plant Gene Resources of Canada (PGRC), including Canada, China, Russia, Germany, Hungary, France, Netherlands, Sweden, South Korea, Japan and Poland were used for the present study (Table 1). Seeds from these varieties were placed in clear Petri dishes lined with two layers of wet filter paper and kept in a growth chamber for a period of 14 to 21 days. They were subjected to a repeating cycle: 16 hours of daylight at 30 °C and 8 hours of darkness at 20 °C. Once the seedlings reached 5 to 10 cm, the roots and seed capsules were removed and the leaves were collected, frozen in liquid nitrogen and stored at -80° until the DNA was extracted.

Table 1: Soybean *Glycine max* accessions used in this study

Accession	Taxon	Name	Origin
CN33248	*Glycine max*	Harosoy63	Canada
CN33251	*G. max*	Harwood	Canada
CN33259	*G. max*	Capital	Canada
CN33275	*G. max*	Maple Arrow	Canada
CN36136	*G. max*	BK17_1_4	Canada
CN39086	*G. max*	X702_3_2	Canada
CN107377	*G. max*	AC Albatros	Canada
CN107380	*G. max*	Medallion	Canada
CN107433	*G. max*	AC Hercule	Canada
CN29744	*G. max*	Seeh Tieh No.5	China
CN29747	*G. max*	Kao Chien Tao	China
CN29791	*G. max*	Feng Shou No.10	China
CN29797	*G. max*	Jin Shen Chi	China
CN30318	*G. max*	Small Golden Yellow No.1	China
CN36008	*G. max*	Gang 7126_9	China
CN43603	*G. max*	Wen Feng 7	China
CN107585	*G. max*	PI358320	China
CN107650	*G. max*	Salut216 China	China
CN107658	*G. max*	Hej He 3	China
CN107467	*G. max*	Grignon 39	France
CN107502	*G. max*	B10	France
CN107504	*G. max*	Grignon 19	France
CN107507	*G. max*	SS	France
CN107509	*G. max*	Tulowka	France
CN107511	*G. max*	Jaune De Desme	France
CN10516	*G. max*	Rouest 13 AI 2	France
CN107517	*G. max*	SEMILUTEA	France
CN107518	*G. max*	Halton	France
CN107515	*G. max*	Geant Vert	France
CN107490	*G. max*	Strain No. 14	Germany
CN107492	*G. max*	Strain No. 42	Germany
CN107497	*G. max*	Strain No. 134	Germany
CN107499	*G. max*	Strain No. 164	Germany
CN107510	*G. max*	Bitterhof	Germany
CN107513	*G. max*	Nordeutsche Swart Matt	Germany
CN107548	*G. max*	Soja_C._St.4/58	Germany
CN107550	*G. max*	Soja_C.St_. 12/58	Germany
CN107561	*G. max*	Soya Heimkraft II	Germany
CN107616	*G. max*	Praemata	Germany

Table1: Continued

Accession	Taxon	Name	Origin
CN107592	*G. max*	Karafuto No. 1	Japan
CN107593	*G. max*	Kamishunbetzu	Japan
CN107595	*G. max*	Shinsei	Japan
CN107625	*G. max*	Ezonishiki	Japan
CN107629	*G. max*	Grignon 48	Japan
CN107630	*G. max*	(Herb 22)	Japan
CN107631	*G. max*	Pulawska Wczesna	Japan
CN107632	*G. max*	A401	Japan
CN107634	*G. max*	B44	Japan
CN107635	*G. max*	Soja 27/60 Heimkraft I	Japan
CN30629	*G. max*	Mica Hungara	Hungary
CN32353	*G. max*	ISZ8	Hungary
CN107557	*G. max*	Keszthelyi Aproszemu Sarga	Hungary
CN107559	*G. max*	Reatz	Hungary
CN107560	*G. max*	Balvanska	Hungary
CN107562	*G. max*	Vince	Hungary
CN107563	*G. max*	Wielnska Brunatna	Hungary
CN107619	*G. max*	PI 378666	Hungary
CN107569	*G. max*	Iregi Nagyszemu Feher	Hungary
CN35309	*G. max*	KAS131_8	South Korea
CN35310	*G. max*	KAS131_9	South Korea
CN35312	*G. max*	KAS133_3	South Korea
CN35313	*G. max*	PGR 7568	South Korea
CN35319	*G. max*	KAS160_2	South Korea
CN35320	*G. max*	PGR 7576	Korea South
CN35344	*G. max*	PGR7640	South Korea
CN35348	*G. max*	KAS581_13	South Korea
CN35352	*G. max*	PGR 7691	South Korea
CN35353	*G. max*	KAS604_23	South Korea
CN107462	*G. max*	No. D.47	Netherland
CN107472	*G. max*	Ras 20	Netherland
CN107475	*G. max*	J_5A	Netherland
CN107481	*G. max*	No.39	Netherland
CN107482	*G. max*	No.47	Netherland
CN107483	*G. max*	No.48	Netherland
CN107484	*G. max*	No.701	Netherland
CN107485	*G. max*	Np.707	Netherland
CN107486	*G. max*	No.709	Netherland
CN10487	*G. max*	No.713	Netherland

Table1: Continued

Accession	Taxon	Name	Origin
CN29403	*G. max*	Amurskaja	Russia
CN30391	*G. max*	Primorskaja	Russia
CN35917	*G. max*	Ussurijskaja	Russia
CN52638	*G. max*	Bisser	Russia
CN52641	*G. max*	Vzlyot	Russia
CN52644	*G. max*	Smena	Russia
CN52645	*G. max*	Seroglazka	Russia
CN107567	*G. max*	Salut216	Russia
CN107572	*G. max*	Urozsajnaja	Russia
CN107547	*G. max*	Bydgoska 052	Poland
CN107552	*G. max*	Zlotka	Poland
CN107553	*G. max*	N. 1954	Poland
CN107554	*G. max*	N. 2054	Poland
CN107637	*G. max*	Bydgoska 057	Poland
CN107638	*G. max*	Bydgoska 071	Poland
CN107639	*G. max*	Bydgoska 074	Poland
CN107640	*G. max*	Czarna Swhn	Poland
CN107642	*G. max*	Zlocista	Poland
CN107643	*G. max*	Zolta Przebeowska	Poland
CN107520	*G. max*	698-1-1	Sweden
CN107525	*G. max*	744-1	Sweden
CN107526	*G. max*	748-5	Sweden
CN107529	*G. max*	749-2	Sweden
CN107533	*G. max*	753-1	Sweden
CN107534	*G. max*	634-13-42-2	Sweden
CN107535	*G. max*	634-20-4-29	Sweden
CN107536	*G. max*	706-4-1	Sweden
CN107537	*G. max*	756-2	Sweden
CN107541	*G. max*	770-3	Sweden

2.2.2 DNA Extraction

Genomic DNA was extracted according to Nkongolo et al., (2001). Frozen leaves were ground into a powder using liquid nitrogen in a mortar and pestle. The samples (between 3g and 5g) were suspended in 15 ml of cetyl trimethylammonium bromide (CTAB) buffer (100 mM Tris HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% (w/v) CTAB) that had been heated to 60 °C and were incubated for 45 minutes. Following the incubation, the aqueous layer was transferred to a new tube and was washed three times with Chloroform:Octonol mix (24:1).

After each wash, the samples were centrifuged for 15 minutes at 17,000 rpm at 25 °C. One volume of the 2-propanol was added to the supernatants in order to precipitate the DNA. The samples were centrifuged for 20 min at a setting of 6,000 rpm at 4 °C. The DNA pellet was desalted in 70% ethanol for 5 min, drained and resuspended in 400 µl of 1XTE (Tris-EDTA).

2.2.3 DNA Quality and Quantification

DNA was quantified using the Fluorescent DNA Quantification Kit (Bio-Rad). DNA quality test was performed to ensure that the DNA was not degraded. Five µl of each DNA sample was run through a 1% agarose gel in 1X Tris-acetate-EDTA (TAE) buffer containing 0.5 µg/ml ethidium bromide. The gels were documented using a Bio-Rad ChemiDoc XRS system and analyzed with Discovery Series Quantity One Software.

2.2.4 DNA Amplification

Forty-four primers (thirty two from ISSR and eleven from RAPD marker system) were chosen for preliminary amplification of DNA from 108 *G. max* varieties (Table 2). All primers were standardized to a 25 ng/ml stock solution. DNA amplification was performed using the procedure described by Nkongolo (1999a, 1999b). PCR (Polymerase Chain Reaction) amplification was performed in a 25 µl volumes that contained 4 mM $MgCl_2$, 2.1 µl of 10x buffer (BioBasics), 200 µM of each dNTP (BioBasics), 0.5 µM primers,10 ng of template and 0.625 units of Taq polymerase (BioBasics). A negative control was included with every reaction. The samples were overlaid with mineral oil and were amplified on an Eppendorf Master Cycler thermal cycler. The thermal cycler was programmed for a 'hot start' of 5 minutes at 95 °C followed by 2 minutes at 85 °C and 42 cycles of 30 sec at 95 °C, 1:30 min at 55 °C and 30 sec at 72 °C, 7 minutes extension at 72 °C.

Table 2: Nucleotide sequence and G+C content for the five ISSR and RAPD primers used to amplify DNA accessions

Primer identification	Nucleotide sequence (5'-3')	G+C content (%)
ISSR 5	ACGACGACGACGGAC	64.28
ISSR Echt 6	ACTCACTCGC	60
UBC 873	GACAGACAGACAGACA	50.00
SC ISSR 6	TTGTTGTTGTTGTTGGB	35.3
ISSR 849	GTGTGTGTGTGTGTGTYA	44.44
RAPD UBC 186	GTGCGTGGCT	70
UBC 337	TCCCGAACCG	70
GRASSE 8	GGGTAACGCC	70
OPA 11	CAATCGCCGT	60
PINUS 23	CCCGCCTTCC	80

The amplified DNA was separated using a 2% agarose gel in 0.5X Tris-Borate-EDTA buffer (TBE) containing 0.5 µg/ml ethidium bromide. The gels were documented using a Bio-Rad Chemidoc XRS system and analyzed with Discovery Series Quantity One 1D Analysis Software.

ISSR and RAPD primers that amplified consistent profiles across the populations were selected for the final analysis.

2.2.5 ISSR and RAPD Analysis

ISSR and RAPD amplification products from each accession from the eleven countries were scored using POPGENE version 1.32 (Yeh et al., 1997). The presence or absence of fragments was scored as 1 or 0 for each band, in order to determine variation between

accessions (within and between countries). The Jaccard's similarity coefficients and genetic distances were generated using FreeTree (Pavlicek et al., 1999).

2.3 Results

2.3.1 DNA Quality and Quantification Analysis

All DNA samples were tested to assess their quality. They all showed a large molecular weight band, indicating that they were not degraded and were deemed suitable for PCR amplification (Figure 1).

2.3.2 ISSR Analysis

Overall, thirty-two primers were screened for amplification of *G. max* DNA (Appendix 1). Five primers generated strong amplification products that were used for genetic analysis (Table 2). The level of polymorphism detected in *G. max* accessions from each country is described in table 4. The highest polymorphic index was observed in accessions from Poland (51.95%) and the lowest in samples from Netherlands (29.22%) (Table 3). Figures 2-8 depict amplified products generated with ISSR primers ISSR 5, ISSR 6 and ISSR 873. ISSR 5 generated the most number of bands (40) followed by ISSR 6 and ISSR Echt 6 with 34 and 33 bands, respectively. ISSR 849 generated the least number of bands (18). The highest level of polymorphic loci was 53. 44% detected with primer ISSR Echt 6. The lowest polymorphism was observed with primer ISSR 5 (Table 4). The genetic distance matrix was generated based on ISSR polymorphism data (Table 5). The accessions from Sweden and Japan were the most genetically closely related with a genetic distance of 0.31. The most genetically distant were the accessions from France and China. Overall, the accessions from different countries were moderately to distantly related (Table 5). In fact, 82% of genetic distance values were above 0.40. The dendrogram revealed two clusters with a low level of confidence (Figure 9). But it clearly revealed with a high level of confidence that the South Korean accessions formed an out-group.

2.3.3 RAPD Analysis

A total of 12 RAPD primers were screened (Appendix 2). Five primers were selected for further study based on their amplification and reproducibility to analyze the DNA samples from all the targeted countries. These primers include OPA 11, Pinus 23, UBC 377, UBC186, and Grasse 8 (Table 2).

Figures 10-15 depict amplified products generated with RAPD primers OPA 11, Grasse 8, Pinus 23, UBC 377 and UBC186. The highest polymorphic index among accessions was 48.00% in France, followed by accessions from Hungary with 43.30%. Over all, the lowest polymorphic index was 29.90%, 26.29% and 24.74% observed in accessions from Canada, Netherlands and China, respectively (Table 6).

At the primer level, primers RAPD 186 and RAPD Grass 8 generated the most number of bands (41) followed by UBC 377 with 40 bands. The lowest number of amplified product was observed with primer Pinus 23. The highest level of polymorphic loci (46.30%) was detected with primer RAPD 186 and the lowest with UBC 377 (Table 7). Overall, there were no significant differences among primers for the detection of polymorphism using all *G. max* accessions.

The genetic distance values based on RAPD data are described in Table 8. Accessions from Canada and Netherlands were the most genetically closely related with accessions from Russia and South Korea being the most distant (Table 8). Overall, RAPD data revealed that the accessions from different countries are closely related with 64% genetic distance values below 0.40. The dendrogram constructed from RAPD data revealed one main cluster. Accessions from Canada and China clustered together with a high degree of confidence.

Likewise, accession from Hungary and Japan also clustered together. Sweden accessions represented an out-group from the rest with a high degree of confidence (Figure 16).

2.3.4 Comparison of ISSR and RAPD Polymorphism

The level of polymorphism generated with ISSR and RAPD primers were compared. Data are summarized in figure 19 and 20. Overall, with the exception of data with Chinese accessions, they were no significant difference between ISSR and RAPD polymorphism data (Figure 17). This was confirmed when the accessions from all the countries were combined to compare ISSR and RAPD data (Figure 18).

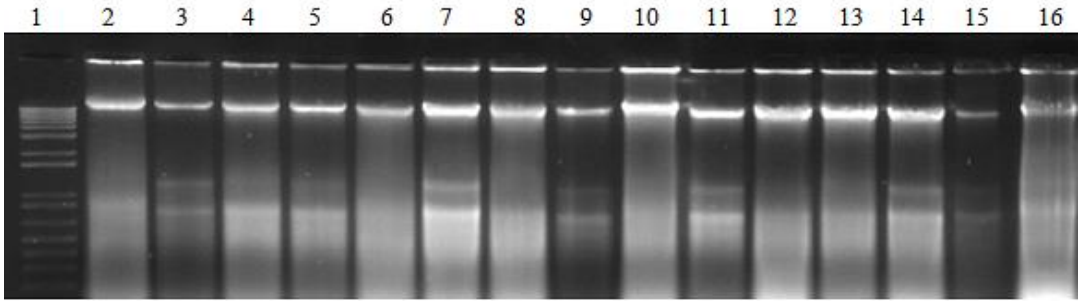


Figure 1: Genomic DNA for quality test of DNA accessions from *G. max*. Lane 1 represents 1Kb^{+} ladder. Lane 2 to 16 contains individual samples of *G. max*

Table 3: Polymorphic loci (%) generated with ISSR primers

Countries	Total number of polymorphic bands	Polymorphic bands(%)
Canada	64	41.56 %
China	57	37.01 %
France	60	38.96%
Germany	57	37.01%
Japan	56	36.36%
Hungary	75	48.70%
South Korea	76	49.35%
Netherlands	45	29.22%
Russia	55	35.71%
Poland	80	51.95%
Sweden	59	38.31%

Table 4: Polymorphic loci (%) generated by each ISSR primer used for each country

	Countries											
Primers	**Canada**	**China**	**France**	**Germany**	**Japan**	**Hungary**	**South Korea**	**Netherlands**	**Russia**	**Poland**	**Sweden**	**Polymorphic bands (%)**
ISSR 5	13/40	12/40	8/40	18/40	14/40	16/40	13/40	12/40	13/40	12/40	10/40	32.05%
ISSR 6	17/35	19/35	18/35	18/35	16/35	19/35	19/35	17/35	15/35	19/35	13/35	49.35%
ISSR Echt6	22/33	19/33	19/33	21/33	13/33	17/33	14/33	18/33	19/33	20/33	12/33	53.44%
ISSR 849	8/18	11/18	8/18	7/18	12/18	8/18	9/18	9/18	8/18	8/18	9/18	48.99%
ISSR873	14/28	11/28	9/28	13/28	11/28	11/28	19/28	8/28	10/28	13/28	13/28	42.87%

Table 5: Distance matrix generated with *Glycine max* ISSR data (FreeTree).

	Countries										
	Canada	**China**	**France**	**Germany**	**Japan**	**Hungary**	**South Korea**	**Netherlands**	**Russia**	**Poland**	**Sweden**
Canada	0.00000	0.60784	0.58252	0.52778	0.50943	0.44915	0.45614	0.50485	0.47619	0.47414	0.36283
China		0.00000	0.73118	0.59804	0.47664	0.46957	0.45133	0.57732	0.53000	0.49558	0.36937
France			0.00000	0.57282	0.49524	0.47368	0.41739	0.58333	0.50495	0.47368	0.38532
Germany				0.00000	0.52885	0.54545	0.43478	0.50980	0.42593	0.44068	0.36607
Japan					0.00000	0.55660	0.39130	0.47525	0.43269	0.44737	0.30973
Hungary						0.00000	0.40164	0.44144	0.42857	0.44262	0.38261
South Korea							0.00000	0.49038	0.43519	0.48696	0.35088
Netherlands								0.00000	0.60000	0.50943	0.41584
Russia									0.00000	0.60000	0.44444
Poland										0.00000	0.48598
Sweden											0.00000

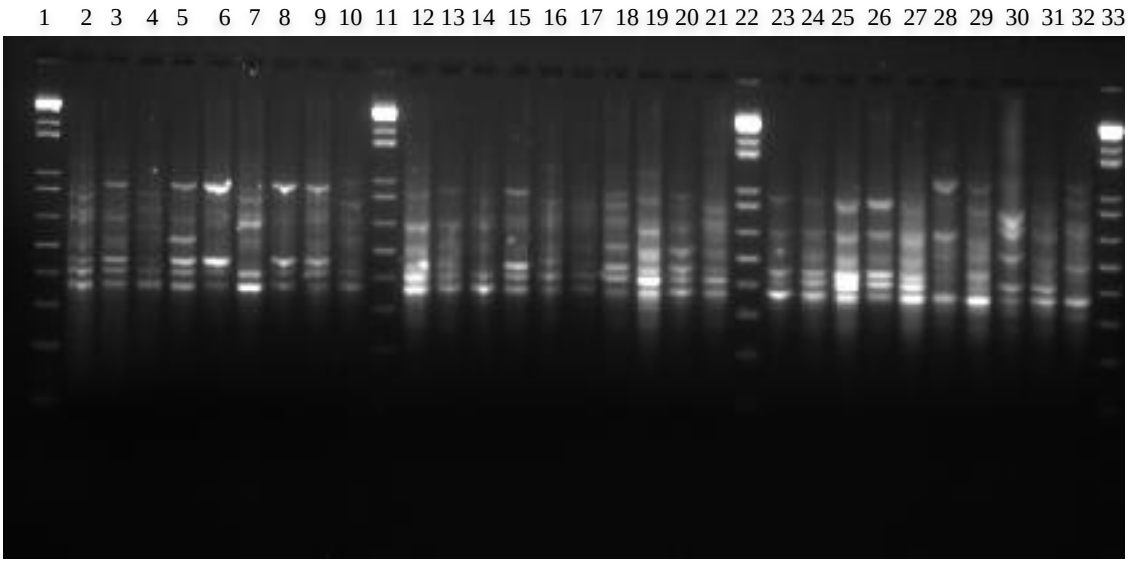


Figure 2: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 5. Lanes 1, 11, 22 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represent accessions from Canada; lanes 12 to 21contains accessions from China and lanes 23 to 32 contains accessions from France.

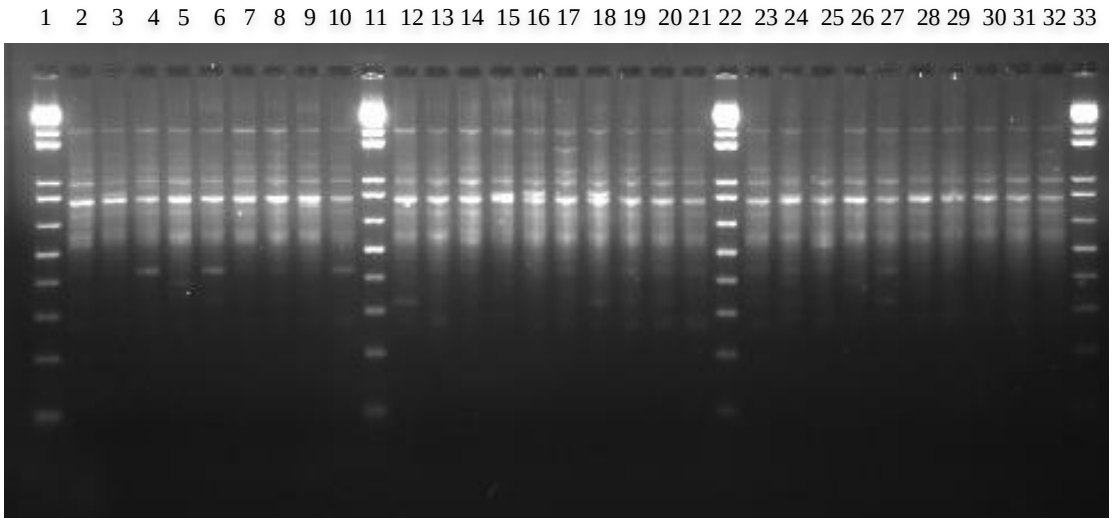


Figure 3: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 6. Lanes 1, 11, 22 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Canada; lanes 12 to 21 contain accessions from China and lanes 23 to 32 contain accessions from France.

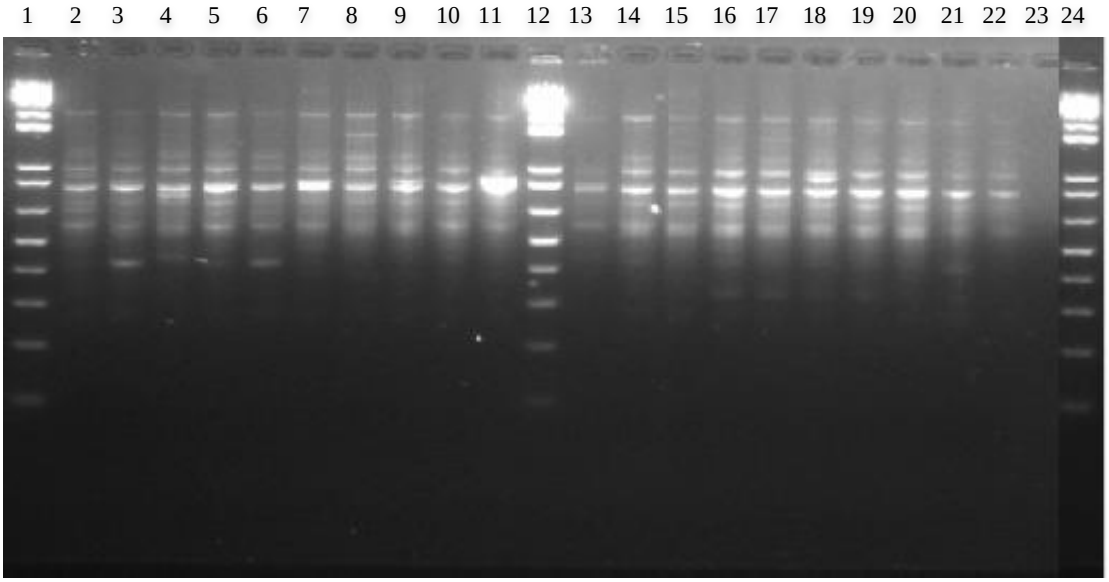


Figure 4: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 6. Lanes 1, 12 and 24 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Germany and lanes 12 to 21 contain accessions from Japan.

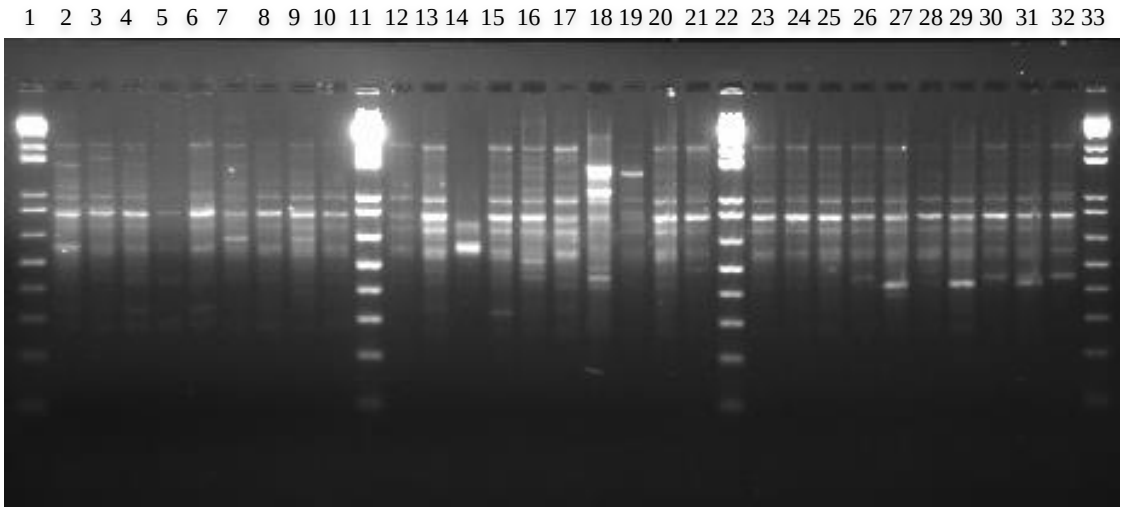


Figure 5: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 6. Lanes 1, 11, 22 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Hungary; lanes 12 to 21 contain accessions from South Korea and lanes 23 to 32 contain accessions from Netherlands.

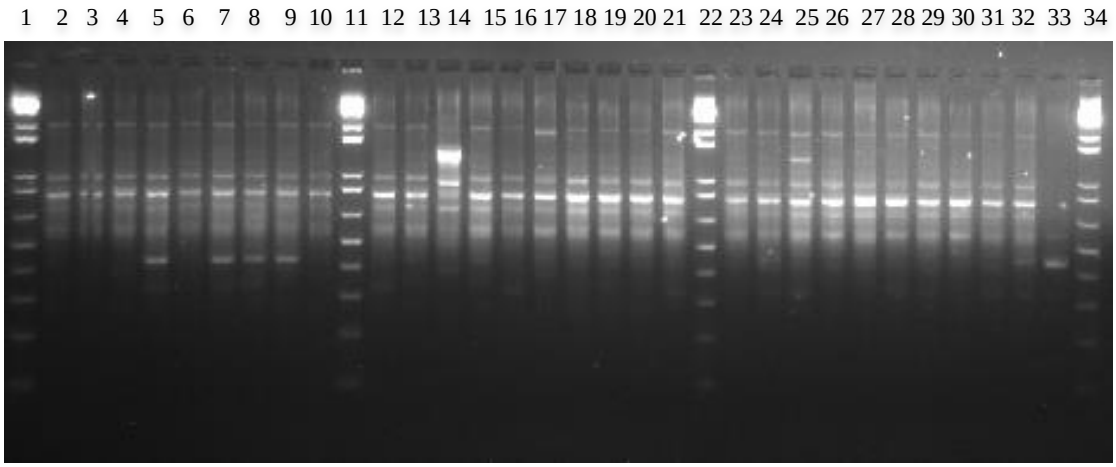


Figure 6: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 6. Lanes 1, 11, 22 and 34 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Russia; lanes 12 to 21 contain accessions from Poland and lanes 23 to 33 contain accessions from Sweden.

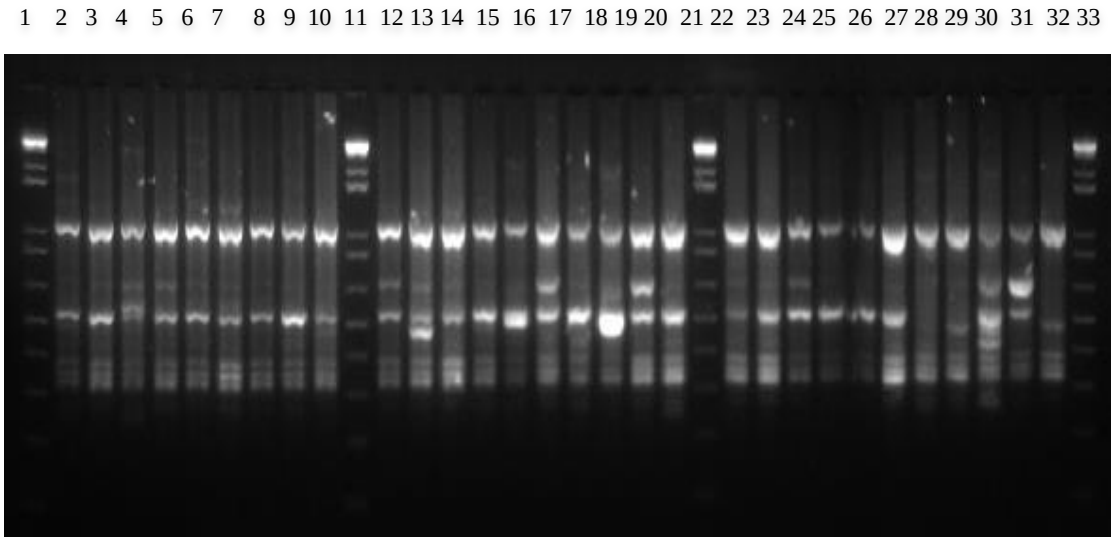


Figure 7: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 873. Lanes 1, 11, 22 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Canada; lanes 12 to 21 contain accessions from China and lanes 23 to 32 contain accessions from France.

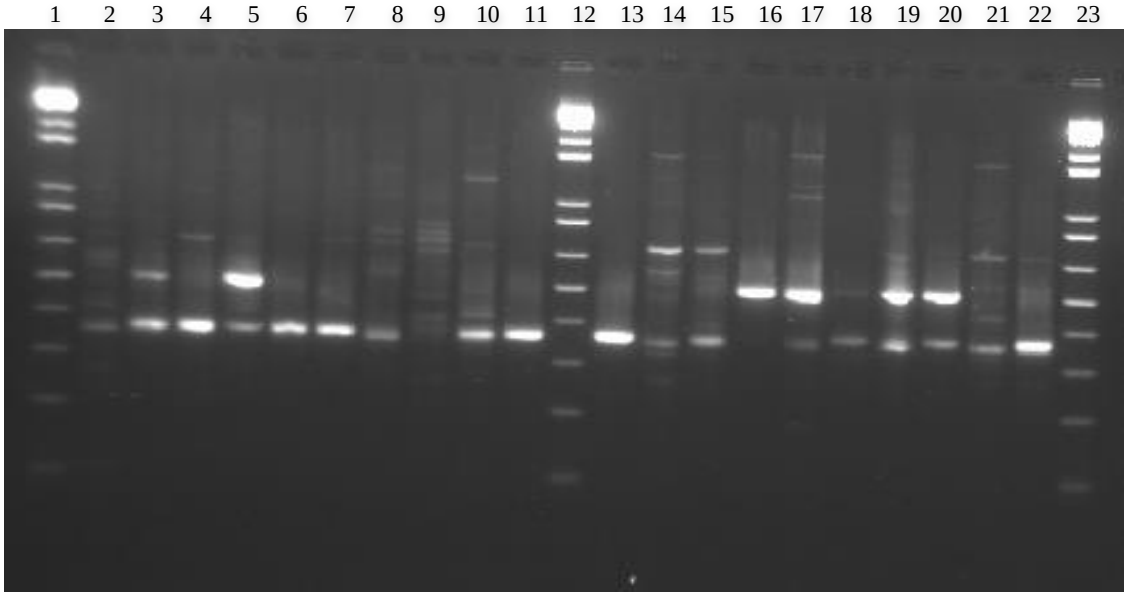


Figure 8: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 873. Lanes 1, 12 and 23 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Poland and lanes 12 to 21 contain accessions from Sweden

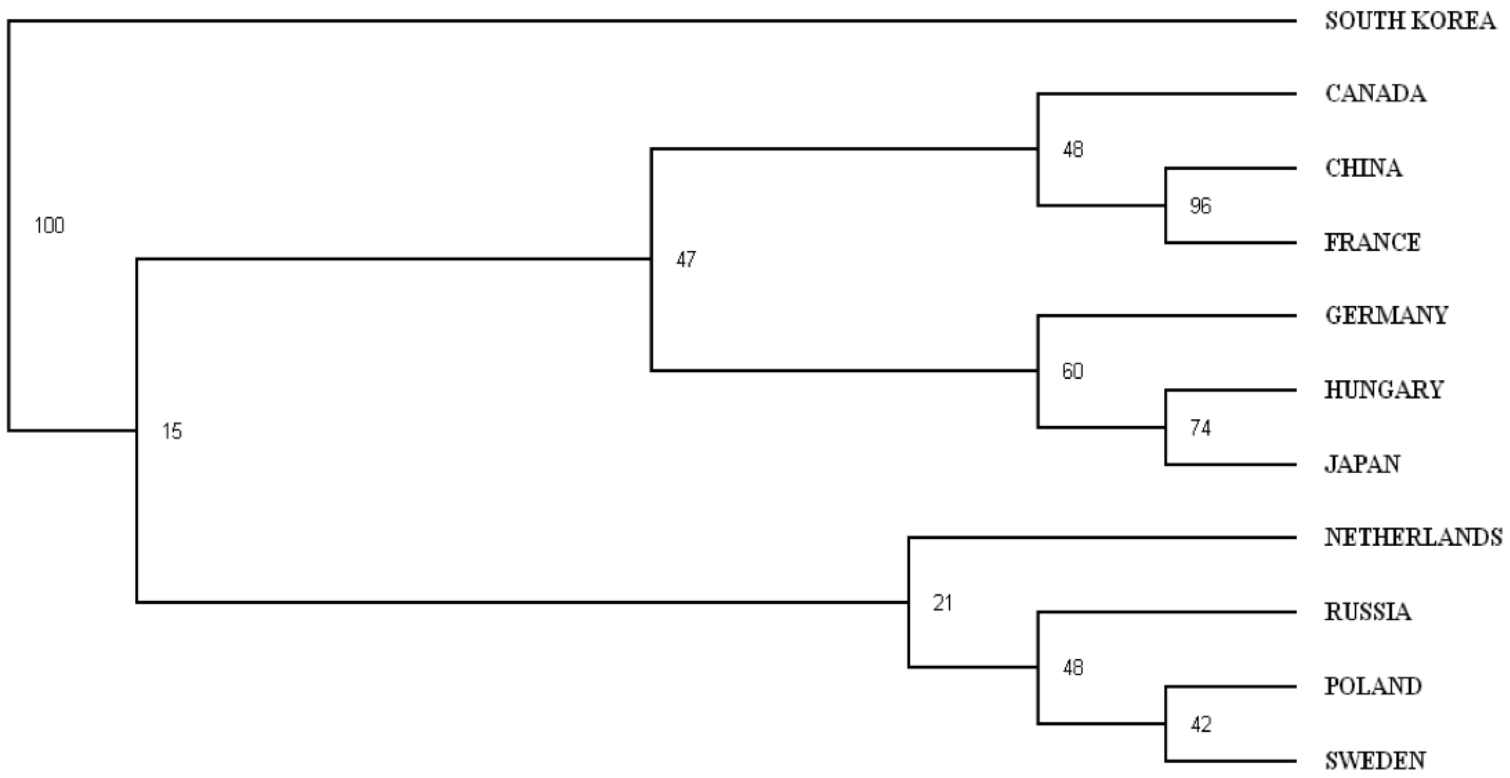


Figure 9: Dendrogram of the genetic relationships between eleven accessions of Soybean (*Glycine max*) from different countries using the data generated from the Jaccard's similarity matrix from ISSR profiles. The values above the branches indicate the patristic distances based on the neighbour- joining (NJ) analysis

Table 6: Polymorphic loci (%) generated with RAPD primers using soybean accessions from different countries.

Countries	Total number of polymorphic bands	Polymorphic bands (%)
Canada	58	29.90%
China	48	24.74%
France	93	47.94%
Germany	69	35.57%
Japan	67	34.54%
Hungary	84	43.30%
South Korea	66	34.02%
Netherlands	51	26.29%
Russia	77	39.69%
Poland	72	37.11%
Sweden	62	31.96%

Table 7: Polymorphic loci (%) generated by each RAPD primer used to amplify DNA from soybean (*Glycine max*) accessions from different countries.

	Countries											
Primers	**Canada**	**China**	**France**	**Germany**	**Japan**	**Hungary**	**South Korea**	**Netherlands**	**Russia**	**Poland**	**Sweden**	**Polymorphic bands (%)**
RAPD UBC 186	10/41	14/41	19/41	19/41	19/41	25/41	20/41	20/41	21/41	24/41	18/41	46.34%
RAPD UBC 377	13/40	8/40	21/40	13/40	18/40	14/40	14/40	11/40	15/40	20/40	16/40	41.28%
RAPD Grasse8	14/41	/1841	21/41	18/41	17/41	20/41	17/41	16/41	18/41	16/41	19/41	43.02%
OPA11	18/38	16/38	22/38	20/38	18/38	18/38	16/38	9/38	21/38	16/38	14/38	44.98%
PINUS 23	22/34	13/34	17/34	13/34	14/34	18/34	14/34	13/34	14/34	14/34	12/34	43.85%

Table 8: Distance matrix generated with *Glycine max* RAPD data (FreeTree).

	Countries										
	Canada	**China**	**France**	**Germany**	**Japan**	**Hungary**	**South Korea**	**Netherlands**	**Russia**	**Poland**	**Sweden**
Canada	0.00000	0.5051	0.40476	0.36752	0.35833	0.35433	0.35656	0.22689	0.32353	0.25564	0.24800
China		0.00000	0.43220	0.43396	0.38393	0.36667	0.34677	0.23214	0.29323	0.27200	0.23333
France			0.00000	0.51240	0.43077	0.42336	0.43478	0.26119	0.39041	0.31034	0.31618
Germany				0.00000	0.40833	0.41270	0.38168	0.25620	0.34783	0.30075	0.28571
Japan					0.00000	0.52101	0.38346	0.29167	0.35000	0.34351	0.30952
Hungary						0.00000	0.41912	0.35537	0.42446	0.38060	0.32824
South Korea							0.00000	0.41525	0.55814	0.43511	0.31111
Netherlands								0.00000	0.47009	0.45872	0.37037
Russia									0.00000	0.50781	0.38931
Poland										0.00000	0.45690
Sweden											0.00000

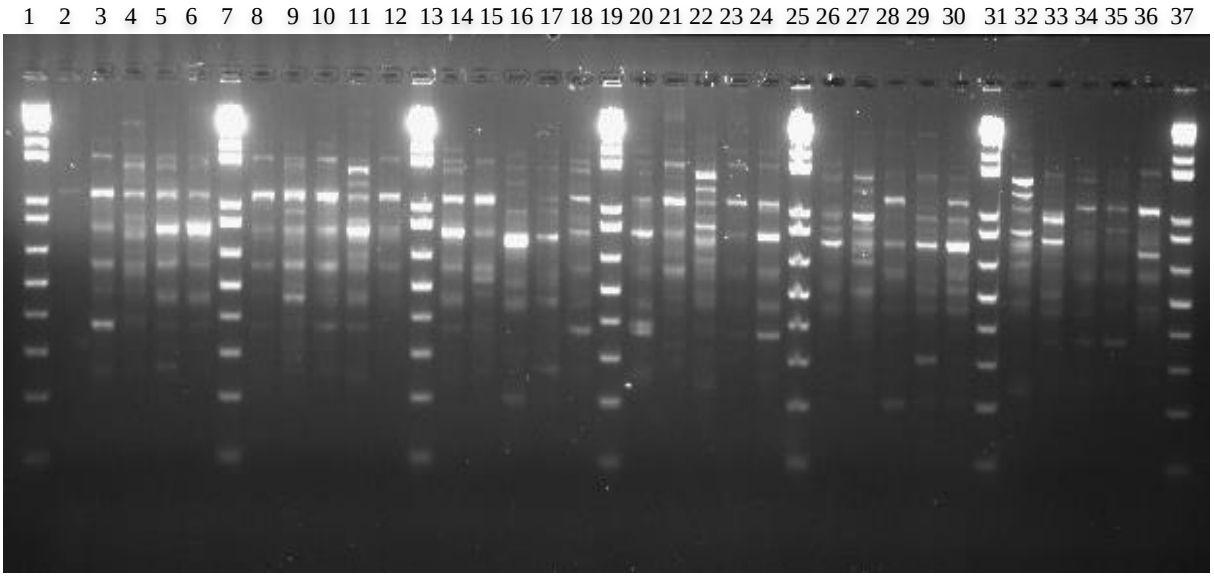


Figure 10: RAPD amplification of Soybean (*Glycine max)* accessions with primer RAPD OPA 11. Lanes 1, 7, 13, 19, 25, 31 and 37 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Canada; lanes 8 to 12 contain accessions from China; lanes 14 to18 contain accessions from France; lanes 20 to 24 contain accessions from Germany; lanes 26 to 30 contain accessions from Japan and lanes 32 to 36 contain accessions from Hungary.

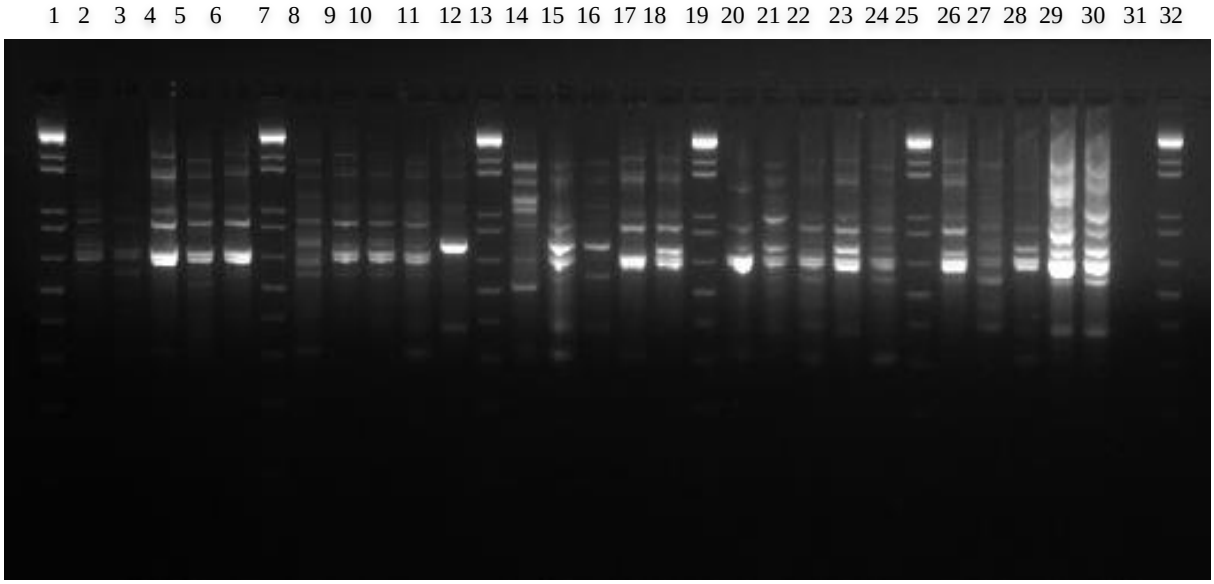


Figure 11: RAPD amplification of Soybean (*Glycine max)* accessions with primer RAPD Grasse 8. Lanes 1, 7, 13, 19, 25 and 32 contains 1Kb^{+} DNA ladder. Lanes 2 to 6 represents accessions from South Korea; lanes 8 to 12 contain Netherlands; Lanes 14 to 18 contains accessions from Russia; lanes 20 to 24 contain accessions from Poland and lanes 26 to 30 contain accessions from Sweden.

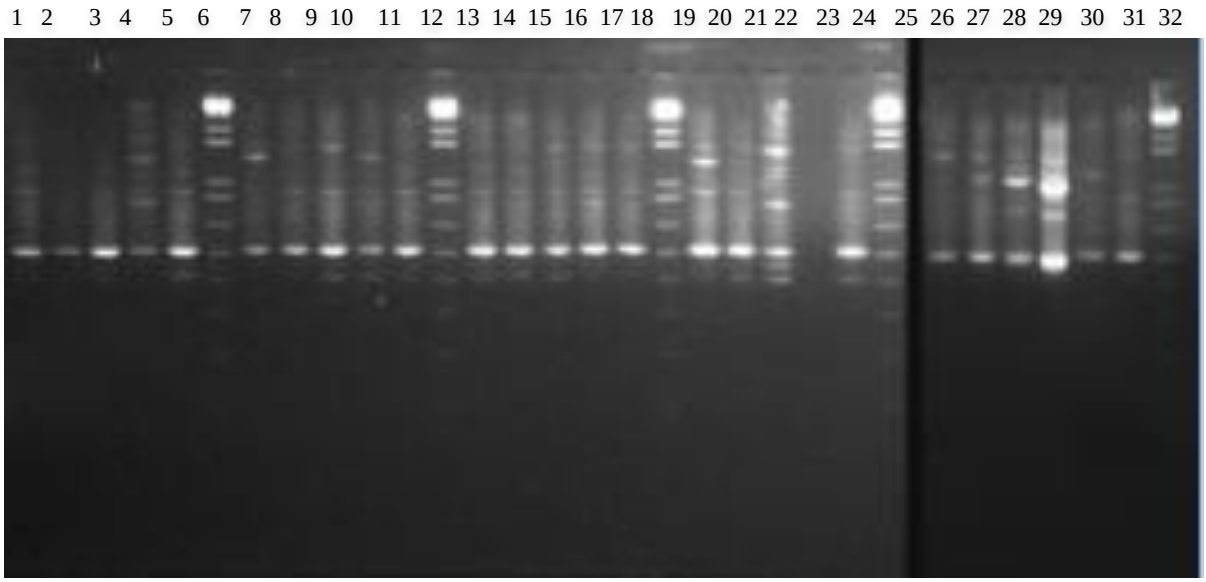


Figure 12: RAPD amplification of Soybean (*Glycine max)* accessions with primer Pinus 23. Lanes 7, 13, 19, 25 and 32 contains 1Kb^{+} DNA ladder. Lanes 1 to 6 represents accessions from South Korea; lanes 8 to 12 contain accessions from Netherland; lanes 14 to 18 contain accessions from Russia; lanes 20 to 24 and 26 contain accessions from Poland and lanes 27 to 31 contain accessions from Sweden.

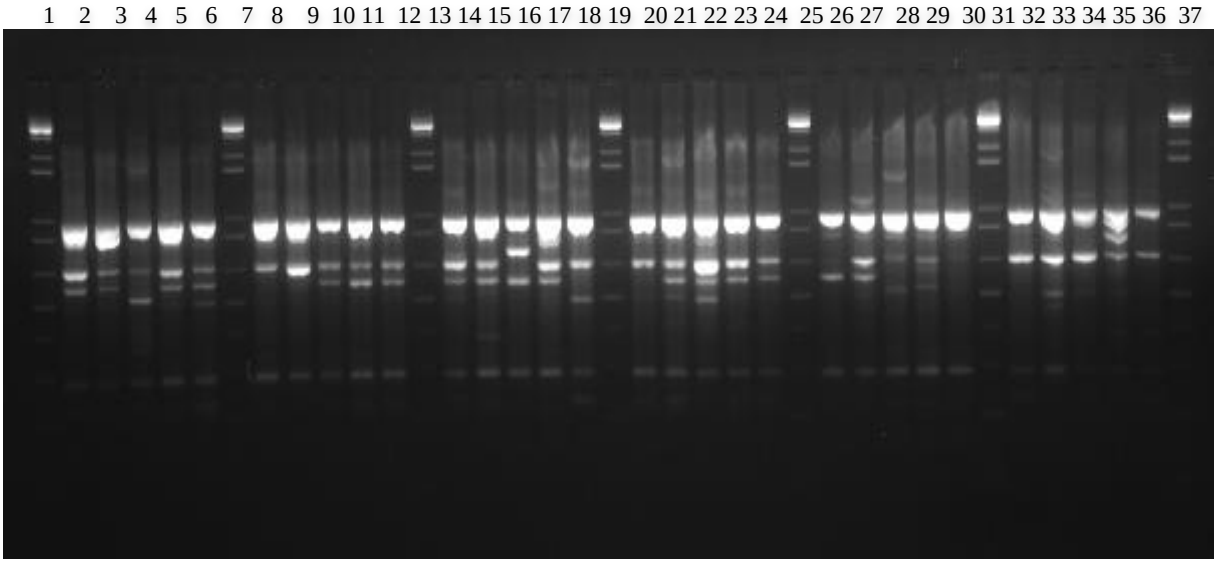


Figure 13: RAPD amplification of Soybean (*Glycine max)* accessions with primer UBC 377. Lanes 1, 7, 13, 19, 25, 31 and 37 contains 1Kb^{+} DNA ladder. Lanes 2 to 6 represents accessions from Canada; lanes 8 to 12 contain accessions from China; lanes 14 to 18 contain accessions from France; lanes 20 to 24 contain accessions from Germany; lanes 26 to 30 contain accessions from Japan and lanes 32 to 36 contain accessions from Hungary.

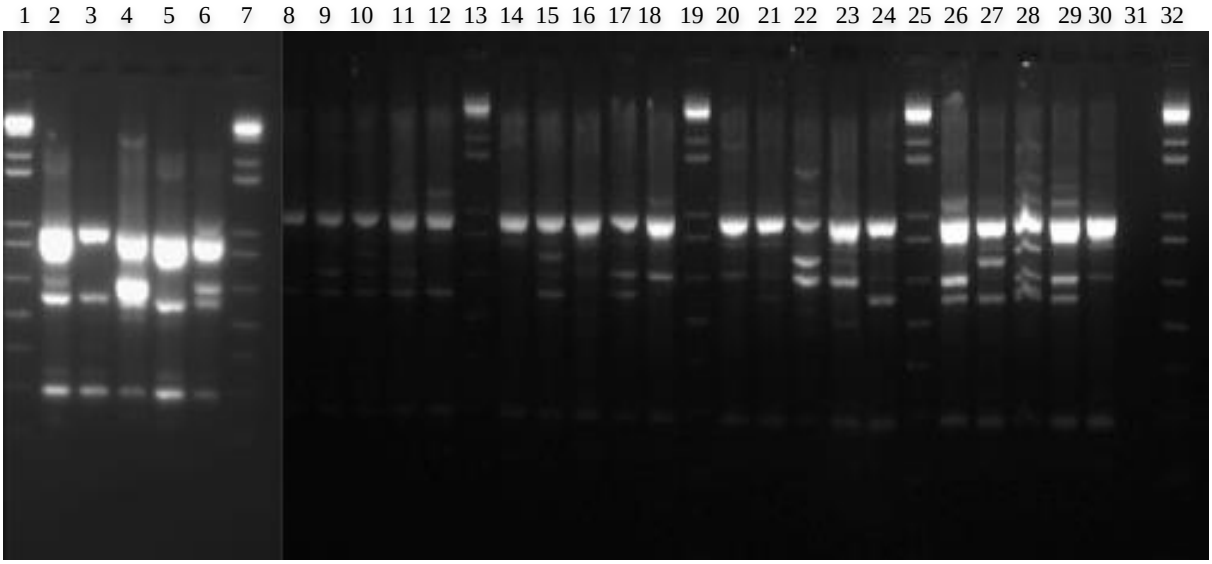


Figure 14: RAPD amplification of Soybean (*Glycine max)* accessions with primer UBC 377. Lanes 1, 7, 13, 19, 25 and 32 contains 1Kb^{+} DNA ladder. Lanes 2 to 6 represent accessions from South Korea; lanes 8 to 12 contain Netherlands; lanes 14 to 18 contain accessions from Russia; lanes 20 to 24 contain accessions from Poland and lanes 26 to 30 contain accessions from Sweden.

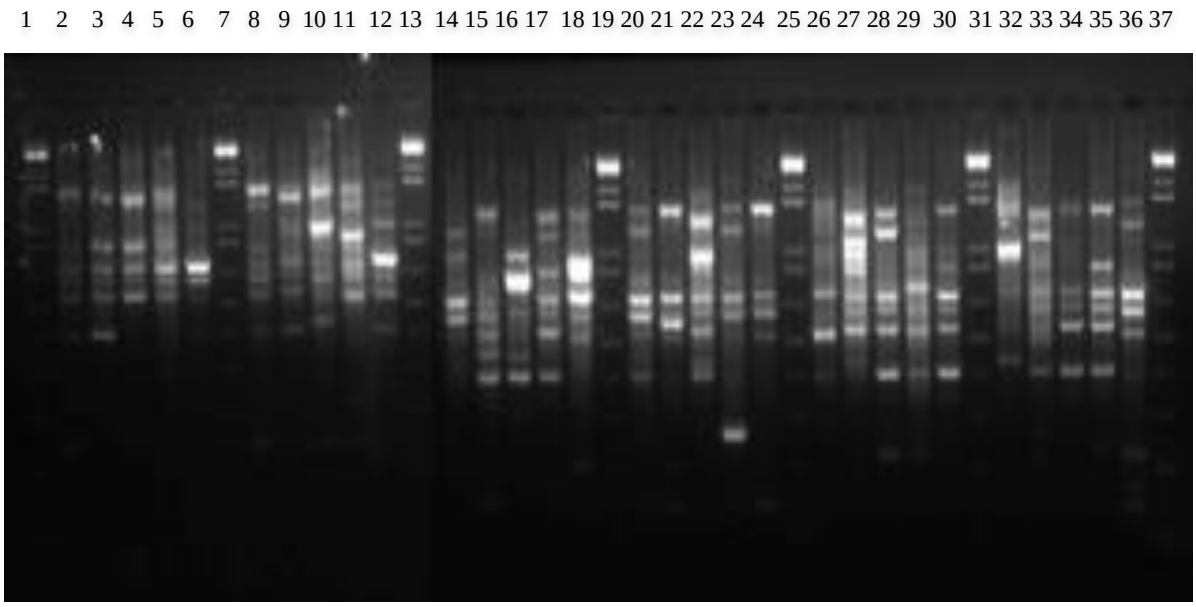


Figure 15: RAPD amplification of Soybean (*Glycine max)* accessions with primer UBC 186. Lanes 7, 13, 19, 25, 31 and 37 contains $1Kb^{+}$ DNA ladder. Lanes 2 to 6 represent accessions from Canada; lanes 8 to 12 contain accessions from China; lanes 14 to 18 contain accessions from France; lanes 20 to 24 contain accessions from Germany; lanes 26 to 30 contain accessions from Japan and lanes 32 to 36 contain accessions from Hungary

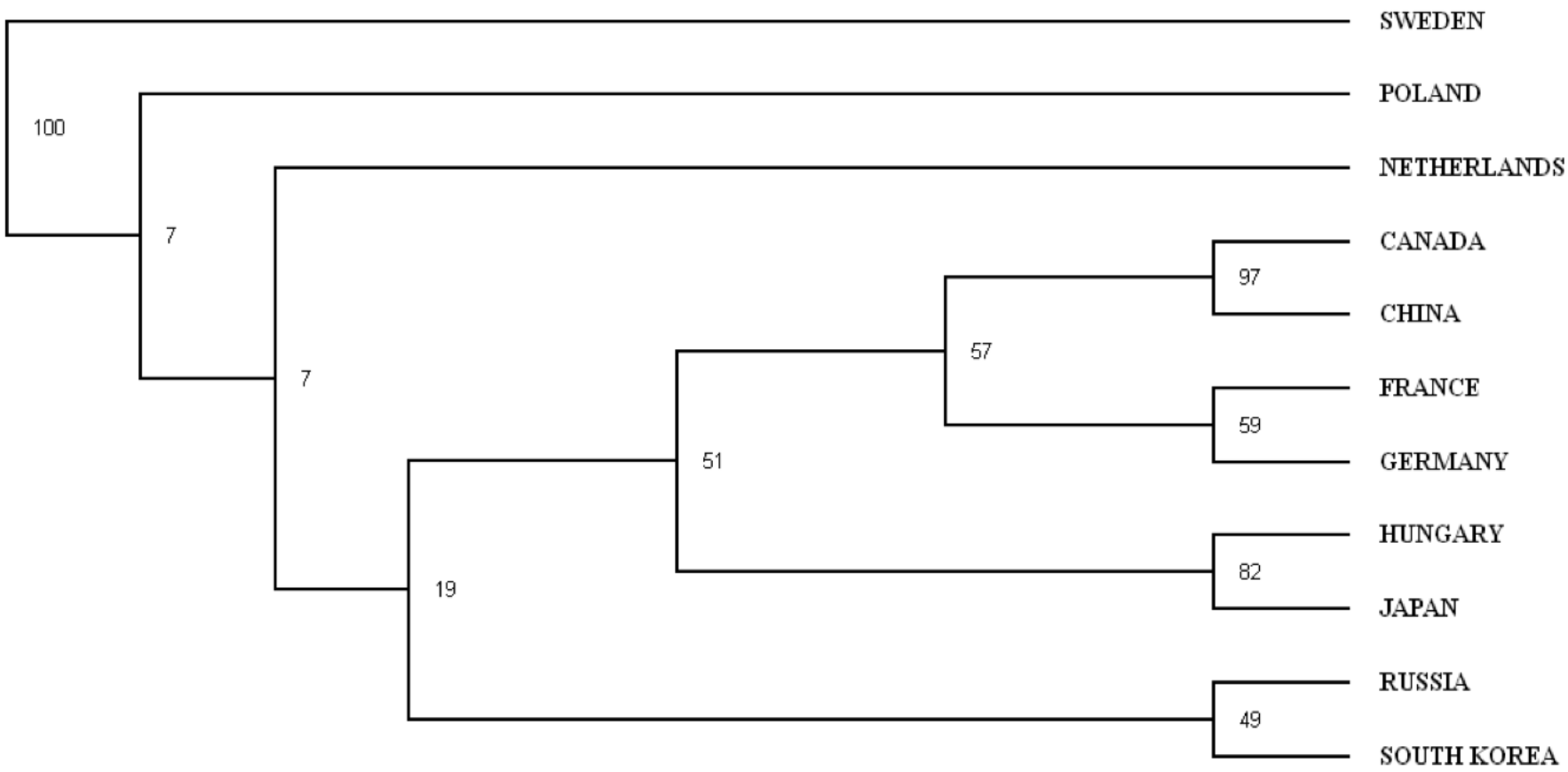


Figure 16: Dendrogram of the genetic relationship between 11 accessions of soybean (*G. max*) from different countries using the data generated from the Jaccard's similarity matrix from RAPD profiles. The values above the branches indicate the patristic distances based on the neighbour-joining (NJ) analysis

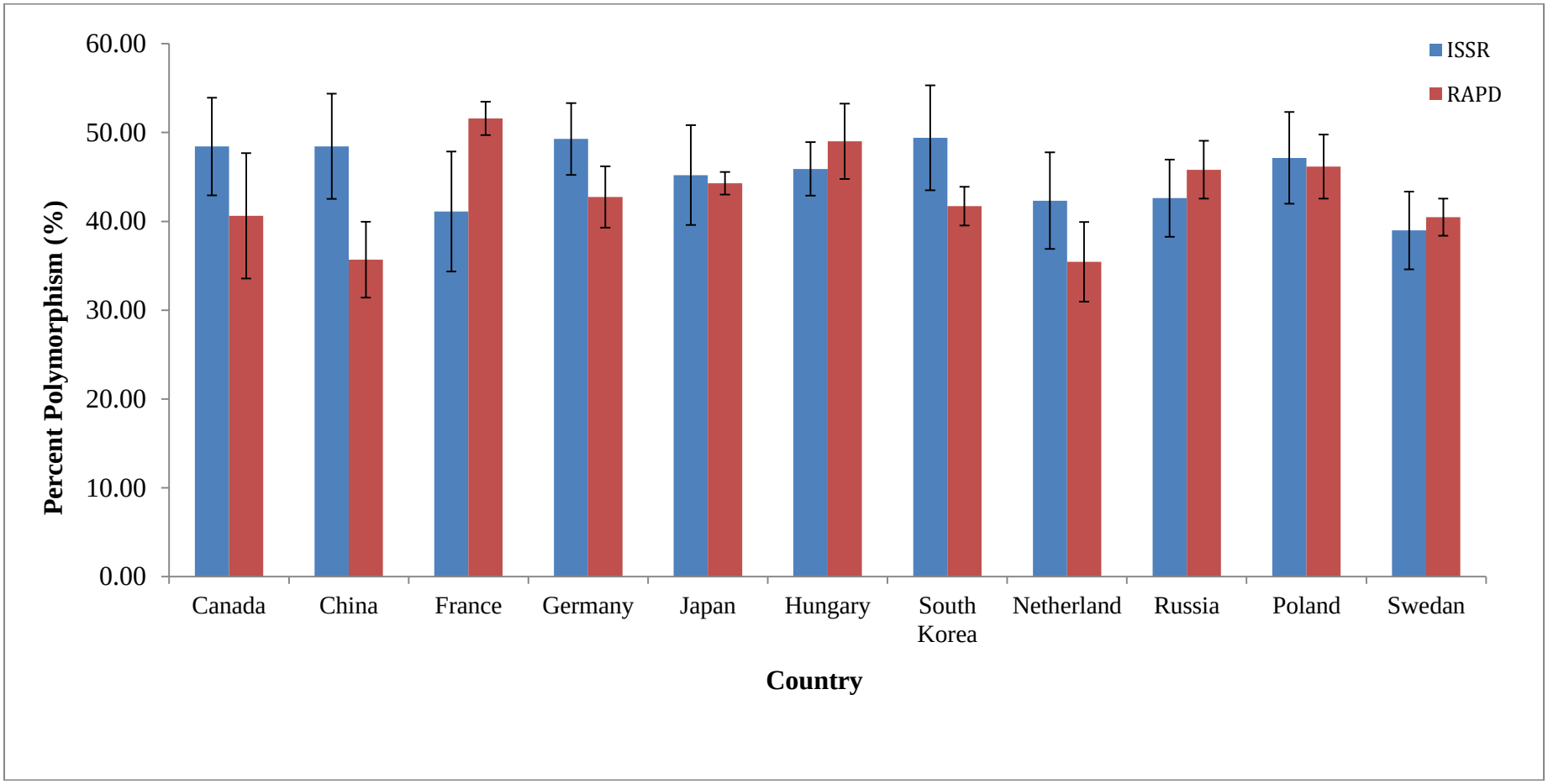


Figure 17: Level of polymorphism generated with ISSR and RAPD primers using soybean (*G. max*) accessions per country.

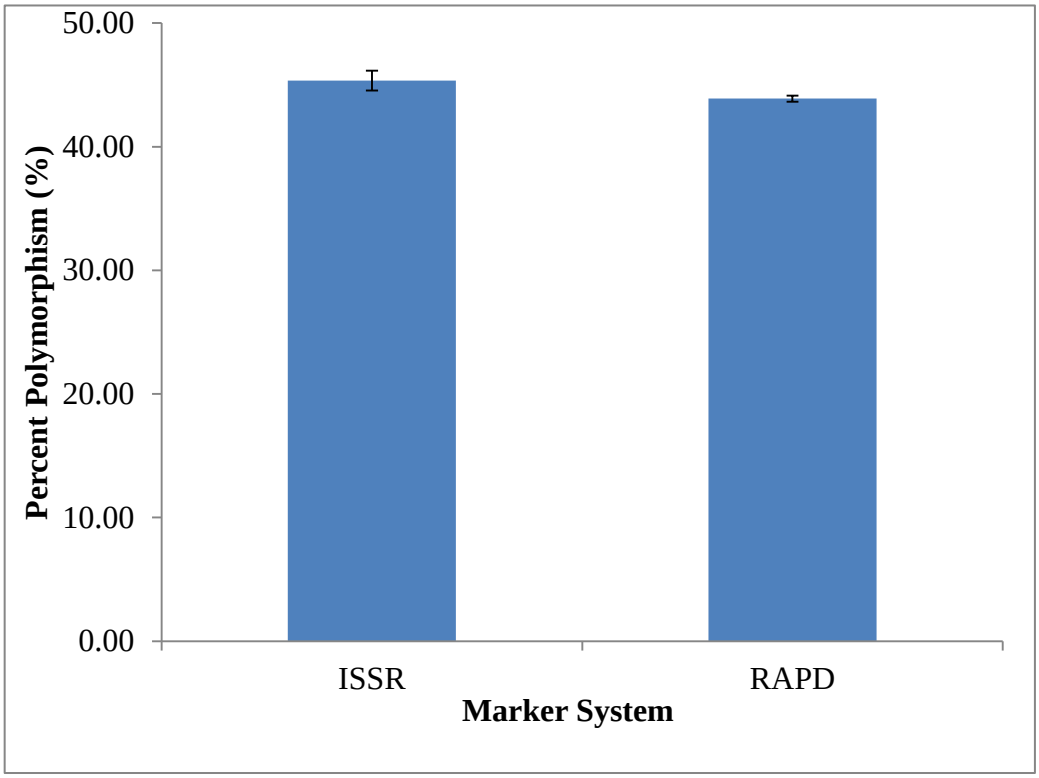


Figure 18: Level of polymorphism generated with ISSR and RAPD primers based on amplification of soybean (*Glycine max*) accessions from different countries

2.4 Discussion

Assessment of genetic diversity in a crop species is a prerequisite to its improvement and helps to generate genetically diversified breeding populations. Considerable variation has been recorded for morphological, physiological and agronomic traits in *G. max* crops (Tshilenge-Lukanda et al., 2012). Various studies have demonstrated that current soybean cultivars are extremely uniform (Hiromoto and Vello, 1986; Brown-Guedira et al., 2000; Priolli et al., 2002; Bonato et al., 2006; Mulato et al., 2010). These studies showed that a few accessions have contributed to the majority of the genes in current cultivars and that the genetic diversity in soybean germplasm is limited. For example as few as five lines account for more than 55% of the genetic background of public cultivars in North America (Brown-Guedira et al., 2000).

In the present study, ISSR and RAPD marker systems were compared in the analysis of the genetic diversity of *G. max*. Overall, the level of polymorphism was similar between ISSR and RAPD based on the analysis of 108 accessions from 11 countries. The highest levels of polymorphisms were 51.5% and 48% for ISSR and RAPD analyzed, respectively. This moderate level of polymorphism is consistent with Chen Y. (2002) who reported 59% of polymorphisms for RAPD analysis of 30 accessions of *G. max*. However, several studies reported low level of polymorphism loci for ISSR and RAPD. For example, Shehzad Faheem (2009) reported 16% of polymorphism in ISSR analysis of *G. max* accessions. Dwivedi et al., (2001) found a level of 18% of polymorphism in *G. max* accessions they analyzed. Baloch et al., (2010) reported 16% of polymorphic loci using ISSR primers. Mudibu et al., (2011) reported 37% of polymorphism loci within the *G. max* genetic pool from the DR-Congo. This low level of genetic variation in *G. max* accessions was validated by Ude et al., (2003) who reported a low level diversity (27%) among 190

Chinese, Japanese and North American soybean cultivars using AFLP (Ude et al., 2003). The ISSR and RAPD data revealed that accessions from Netherlands and China are the least genetically variable. Low levels of genetic variation in *G. max* have been attributed to the small number of parental accessions originating from the same source (Ude et al., 2003).

Overall, ISSR and RAPD primers used in the present study revealed a similar level of polymorphism for the *G. max* accessions analyzed. Previous studies have shown different levels of polymorphism within and among different varieties and species when RAPD and ISSR were compared. For example, Fang and Roose (1997) reported high levels of interspecific variation with RAPD markers than with ISSR markers in Citrus species. On the other hands, several authors detected high level of polymorphism with ISSR system compared to RAPD in several plants (Nkongolo et al., 2005; Raina et al., 2001; Rus-Kortekaas et al., 1994; Nagaoka and Ogihara 1997; Qian et al., 2001).

Technically, RAPD and ISSR markers target different areas in the genome. RAPD markers reveal polymorphisms in coding and non-coding regions, as well as repeated or single copy sequences covering the entire genome (Williams et al., 1990). The origin of the ISSR amplification products is known to be from the sequences between two simple-sequence repeat (also known as Microsatellite) primer sites where length variation does not necessarily reflect simple-sequence length polymorphism (Zietkiewicz et al., 1994). Microsatellite loci are dispersed throughout the genome and are hypervariable because of DNA slippage (Semagn et al., 2006). Most often ISSR detects more polymorphisms than RAPD primers because of the high levels of variability in microsatellite loci. The discrepancy between variations revealed by RAPD and ISSR result from different targeted genomic areas, which undergo a different evolutionary process due to selection forces

(Qian et al., 2001). Different genetic information is generated when RAPD and ISSR molecular marker techniques are used to assess the inter-specific and intra-specific variability. The level of variation detected with each system greatly depends on the primer used therefore making comparisons for the levels of polymorphism generated with ISSR and RAPD marker systems inappropriate.

Among other molecular marker systems, microsatellites and AFLP have been widely used to assess genetic diversity among and within populations (Zhang et al., 2006; Prasad et al., 2000; Rus-Kortekass et al., 1994). Microsatellites amplification is expected to produce a single marker since the microsatellite primers target a single locus (Prasad et al., 2000). This means that many reactions would be needed to properly determine a population's genetic variability. AFLP has been gaining popularity over microsatellites in these types of studies. This is in part because AFLP assay is equivalent to the amplification of several microsatellite primers (Roy et al., 2004). But, ISSR and RAPD primers are easier to use than AFLP and microsatellites (Arslan and Okumus, 2006; Mattioi et al., 2002; Matos et al., 2001; Gibert et al., 1999; Nagaoka and Ogihara 1997). In conclusion, these two systems (ISSR and RAPD) should be used conjointly to assess genetic diversity because they are targeting different regions of the *G. max* genome.

Chapter 3: Identification and Characterization of Genome-Diagnostic and Specific ISSR, RAPD and SCAR Markers.

3.1 Introduction

Breeders have successfully increased the genetic variability in the *G. max* gene pool (Tumanova et al., 2008). Molecular markers linked to major quantitative trait loci (QTL) are useful for marker-assisted selection in soybean-breeding programs (Funatsuki et al., 2005; Panthee et al., 2006). They facilitate the estimation of the diversity degree and genetic constitution of the existing germplasm, as well as, the prediction of the heterotic effects based on the genetic distance between the parents in hybrid programs, contributing to soybean breeding efficiency (Doldi et al., 1997; Narvel et al., 2000; Sudaric et al., 2008; Mladenovic et al., 2008). Soybean breeding can be divided into two major categories: molecular genetics and genetic transformation. Molecular genetics studies indicate how genetic information is encoded within the DNA and how biochemical processes of the cell translate the genetic information into the phenotype. (Sudaric et al., 2004; Shoemaker et al., 2004). The present study reports for the first time the development and characterization of accession-diagnostic marker in several soybean breeding programs.

3.2 Materials and Methods

3.2.1 Identification of Variety-Diagnostic ISSR and RAPD Markers

All genomic DNA extracted was amplified with ISSR and RAPD primers as described in chapter 2. The ISSR and RAPD bands that were present in one variety and absent in all others were classified as variety-diagnostic markers (Table 10 and 11).

3.2.2 Cloning and Sequencing

ISSR variety-diagnostic marker was extracted from agarose gel using QIAquick Gel Extraction Kit (Qiagen). DNA fragments were excised from the gel and were quantified using the Spot Test method (Sambrook et al., 1989). The selected DNA marker was cloned using the pGEM-T Easy Vector System II (Promega). The DNA fragment was ligated into the multiple cloning site of the pGEM-T Easy Vector (Figure 19).

For optimal results, an insert:vector ratio 1:1, 1:2 and 1:3 were used in the ligation reaction. All ligation reactions were incubated overnight at 4 °C. The recombinant plasmids containing either a DNA fragment extracted from the gel or a positive control insert, as well as background control were added to *Eschcericia coli* JM109 high efficiency competent cells (Promega). The cells were heat-shocked for 45 seconds at 42 °C in water bath and then transferred to an ice bath for 2 minutes. SOC medium was added and incubated at 37 °C at 150 rpm for 1.5 hours. One hundred microliters of each culture was transferred to LB/ampicillin/IPTG/X-Gal plates using a bent glass rod. These plates were then incubated overnight at 37 °C and then at 4 °C for an hour. White colonies were randomly picked with an inoculating loop and were added to 5 ml of LB/ampicillin broth in a lightly capped 15 ml disposable centrifuge tube. These tubes were incubated at 37 °C at 200 rpm overnight. The plasmid DNA was purified using QIAprep Spin Miniprep Kit (Qiagen). The purified recombinant plasmids were sent to NAPS unit (University of British Columbia) and sequenced using an ABI 373 auto sequencer (PE Biosystem) with four fluorescent dyes. Computer assisted analysis of the sequence data was performed using ABI Prism sequencing analysis version 3.3 (PE Biosystem).

3.2.3 Sequenced Characterized Amplified Region (SCAR) markers

Primers flanking the insert were designed and synthesized based on the sequences from the cloned ISS sequences. These primers were optimized to ensure their specificity and to produce a clear banding pattern and were used for PCR amplification of *G. max* samples following the same protocol as previously described in chapter 2.

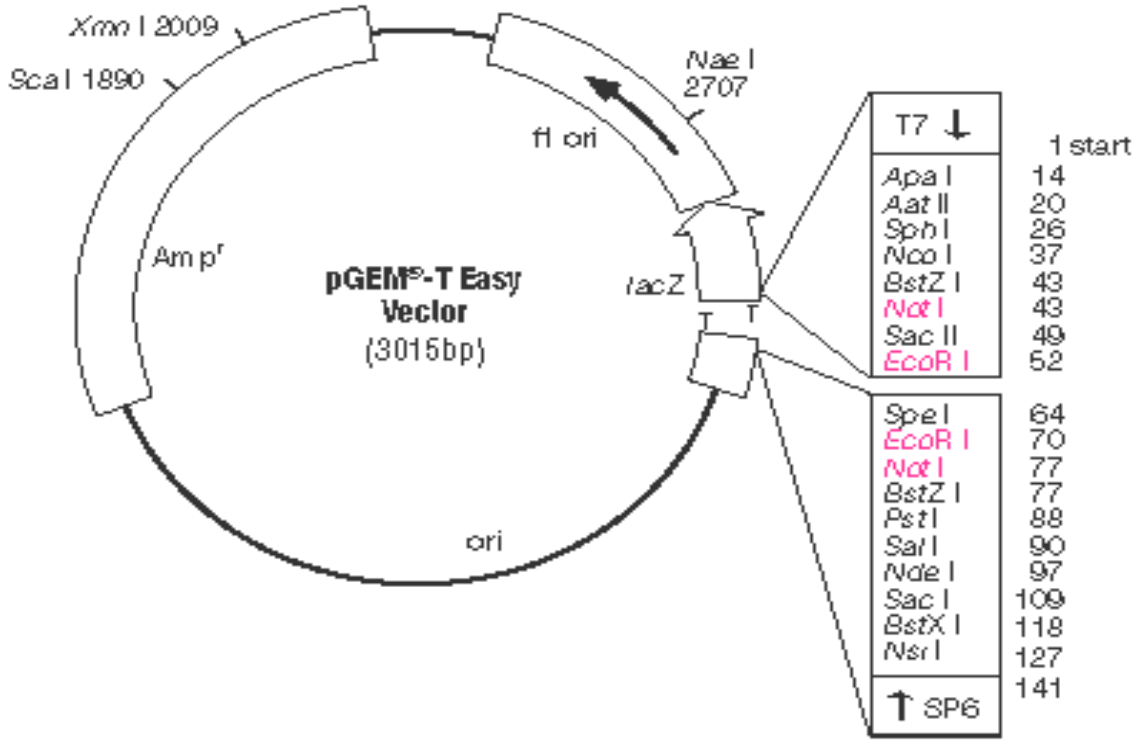


Figure 19: Map of the pGEM-T Easy Vector and reference points used in this study

3.3 Results

Screening of all ISSR and RAPD primers were performed using 108 varieties of *G. max* accessions from eleven countries. This essay was repeated twice to validate the results.

3.3.1 Variety-Diagnostic ISSR Marker

Of the thirty-two ISSR primers screened, five produced several variety-diagnostic markers between countries accessions described in table 9. Primer ISSR 5 generated several variety-diagnostic markers with DNA samples from other countries (Table 9). Other variety-diagnostic markers generated with different ISSR primers are described in table 9. In general for ISSR markers, 19 variety-diagnostic markers were identified in accessions from Canada, 11 from France, 18 from China, 24 from Germany, 14 from Japan, 18 from Hungary, 20 from South Korea, 14 from Netherlands, 15 from Russia, 23 from Poland and 15 from Sweden (Table 9). Appendices 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 and 16 depict amplified product using different primers and show some variety-diagnostic markers. More importantly, ISSR 5 primer generated a variety-diagnostic marker unique to accession *Kao Chien Tao* from China relative to all other accessions screened. This was the only one of the five ISSR primers screened that generated variety-diagnostic band across all the accessions.

3.3.2 Variety-Diagnostic RAPD Marker

Of the twelve RAPD primers screened, five primers produced variety-diagnostic markers. The amplification of the genomic DNA with RAPD UBC 186 generated 84 variety-diagnostic markers with samples from 11 countries. Details of all the variety-diagnostic markers that were identified are described in table 10. In general for RAPD markers we identified 17 variety-diagnostic markers in accessions from Canada, 42 from France, 19

from China, 26 from Germany, 30 from Japan, 34 from Hungary, 19 from South Korea, 16 from Netherlands, 36 from Russia, 33 from Poland and 30 from Sweden (Table 10). Appendix 12, 13 and 14 depict amplified product using different primers and show some variety-diagnostic markers.

3.3.3 SCAR Marker Analysis

Attempt to develop a SCAR marker were made only for the 303 bp ISSR marker that was diagnostic for the *Kao Chien Tao* accession from China (Figure 20). To this end, this 303 bp band was successfully cloned into *E.coli* JM109 and then subsequently sequenced (Figure 21). The consensus sequence of the variety-diagnostic marker is presented in figure 22. A pair of primers targeting the inserts was designed to produce a SCAR marker. Designed primers named SCAR 4 (F) and SCAR 4 (R) amplified a band at 230 bp from the control plasmid (Table 11, Figure 22). The amplification of DNA with this pair of primers using all the genomic DNA to confirm the specificity of the SCAR marker developed did not produce any amplified products.

Table 9: Variety-diagnostic ISSR markers for soybean (*Glycine max*) accessions within each country

Primers	Accession	Marker size (bp)
	Canada	
ISSR5	*Harosoy63*	312, 760, 980, 1137
	Capital	587
ISSR849	*Harosoy63*	1687
ISSR Echt6	*Capital*	400
	Harwood	394, 640
	Harosoy63	723, 934
	BK17_1_4	950, 1436
ISSR6	*BK17_1_4*	334
	Maple Arrow	374
	Harwood	609
ISSR873	*BK17_1_4*	808
	Harosoy63	671, 1714
	China	
ISSR5	*Kao Chien Tao*	558
	Seeh Tieh No.5	587
	Small Golden Yello No.1	1074
ISSR849	*Small Golden Yello No.1*	340
	Jin Shen Chi	406
	Seeh Tieh No.5	460
	Kao Chien Tao	1107
ISSR Echt6	*Small Golden Yello No.1*	520
	Jin Shen Chi	800
	Seeh Tieh No.5	1020
	Small Golden Yello No.1	1050
	Kao Chien Tao	934, 1271
ISSR6	*Small Golden Yello No.1*	875, 1570
ISSR873	*Kao Chien Tao*	475
	Jin Shen Chi	482
	Small Golden Yello No.1	600
	France	
ISSR5	*Tulowka*	520, 664
ISSR Echt6	*B10*	350
	Grignon 39	400
	Grignon 19	950, 1436
	SS	1970
ISSR6	*B10*	875, 1040
ISSR873	*Grignon 39*	875, 1040

Table 9: Continued

Primers	Accession	Marker size (bp)
	Germany	
ISSR5	*Strain No. 134*	260
	Strain No. 42	587
	Strain No. 14	466, 723, 800,
	Strain No. 164	760
	Bitterhof	500, 558, 685, 783
ISSR849	*Strain No. 164*	460
ISSR Echt6	*Strain No. 134*	520
	Bitterhof	1436
	Strain No. 42	400, 640
	Strain No. 164	950, 1271
ISSR6	*Bitterhof*	285
	Strain No. 134	445, 766, 1761
ISSR873	*Strain No. 164*	197
	Strain No. 42	600, 1343
	Japan	
ISSR5	*Grignon 48*	403
	Kamishunbetzu	342, 466, 500, 1250
ISSR849	*Kamishunbetzu*	340
	Grignon 48	900
ISSR Echt6	*Kamishunbetzu*	750, 934, 1530
ISSR6	*Karafuto No. 1*	875
	Grignon 48	1454
	Shinsei	1761
ISSR873	*Grignon 48*	752
	Hungary	
ISSR5	*ISZ8*	260 ,390
	Balvanska	500
	Mica Hungara	466 ,760 ,783
	Reatz	1000
ISSR849	*Reatz CN107559*	460
ISSR Echt6	*Balvanska*	430
	Reatz	350 ,520 ,1020
	Mica Hungara	1436
ISSR6	*Mica Hungara CN30629*	400
	Balvanska 107560	875
	ISZ8 CN32353	1671
ISSR873	*Reatz CN107559*	1632
	Balvanska CN107560	1923

Table 9: Continued

Primers	Accessions	Marker size (bp)
	South Korea	
ISSR5	*KAS160_2*	520,1450
	PGR 7568	760
	KAS133_3	1137
ISSR849	*KAS160_2*	340
	KAS131_9	406
	KAS133_3	1687
ISSR Echt6	*KAS160_2*	1436
ISSR6	*KAS160_2*	530
	KAS133_3	609
ISSR873	*KAS160_2*	290
	KAS131_8	368
	KAS133_3	197, 387
	PGR 7568	414, 432, 450, 600, 650,752
	Netherlands	
ISSR5	*No.39*	1430
ISSR849	*No.47 CN107482*	340
	J_5A CN107475	626
	No. D.47 CN107462	1107
ISSR Echt6	*Ras 20*	320
	No.47	350
	J_5A	800
	No.39	400, 1020, 1050, 1530
ISSR6	*No.47*	374
	No.39	474
ISSR873	*J_5A*	521
	Russia	
ISSR5	*Amurskaja*	611
	Ussurijskaja	723
	Vzlyot	1450
	Primorskaja	1954
ISSR849	*Vzlyot*	440
ISSR Echt6	*Vzlyot*	1100, 1271
	Amurskaja	1530, 1665
	Bisser	400, 1922
ISSR6	*Bisser*	334
	Ussurijskaja	875
ISSR873	*Amurskaja*	521
	Primorskaja	1237

Table 9: Continued

Primers	Accession	Marker size (bp)
	Poland	
ISSR5	*Bydgoska 057*	611, 1220
	Zlotka	940
	N. 2054	1250
	N. 1954	1630
ISSR849	*Bydgoska 052*	460
ISSR Echt6	*N. 1954*	400, 453
	N. 2054	723
	Bydgoska 057	820
	Zlotka	1190
	Bydgoska 052	1271
ISSR6	*N. 2054*	272
	N. 1954	374, 445, 850, 927
	Bydgoska 052	1040
ISSR873	*Bydgoska 052*	241, 352, 482, 922, 1343
	Sweden	
ISSR5	*753-1*	1250
	748-5	1324
	749-2	685, 1738
ISSR849	*753-1*	875
	748-5	1107
ISSR Echt6	*748-5*	950
	744-1	1020
	698-1-1	1436
ISSR6	*698-1-1*	766
	753-1	850
	748-5	1349
ISSR873	*744-1*	316, 600
	753-1	922

Table 10: Variety-diagnostic RAPD markers for soybean (*Glycine max*) accessions within each country.

Primers	Accession	Marker size (bp)
	Canada	
OPA11	*Harosoy63*	868, 1769
	Capital	323, 385, 538, 650, 1224
RAPD 377	*Capital*	261, 413, 434, 469
UBC186	*Maple Arrow*	439
	Harosoy63	793
RAPD Grass 8	*Harosoy63*	706
	Harwood	900
Pinus 23	*Maple Arrow*	830, 10000
	China	
OPA11	*Kao Chien Tao*	385, 949
	Jin Shen Chi	622, 868, 1274
UBC186	*Feng Shou No.10*	438, 521
	Jin Shen Ch	849
	Kao Chien Tao	1404
RAPD Grass 8	*Jin Shen Chi*	608
	Seeh Tieh No.5	671, 706, 1026
	Feng Shou No.10	1007
Pinus 23	*Seeh Tieh No.5*	455, 1044
	Jin Shen Chi	660
	Small Golden Yello No.1	630, 1451
	France	
OPA11	*Grignon 19*	200, 526,1768
	B10	385, 483, 538
	Tulowka	895, 930
	SS	265
	Grignon 39	2000
RAPD 377	*B10*	238, 285, 413
	Tulowka	434, 469, 1096
	Grignon 39	531
	SS	627, 850, 936
	Grignon 19	551, 869
UBC186	*B10*	134, 1552
	Grignon 19	311
	Tulowka	206, 521, 966
RAPD Grass 8	*Tulowka*	352
	Grignon 39	503, 624
	Grignon 19	396, 474, 816, 863, 948
Pinus 23	*Tulowka*	270, 560, 859, 1145
	SS	630, 660

Table 10: Continued

Primers	Accession	Marker size (bp)
	Germany	
OPA11	*Strain No. 14*	265, 1500
	Strain No. 42	690, 1190
	Strain No. 134	650, 769, 868, 1274
RAPD 377	*Strain No. 42*	238, 1096
	Strain No. 134	850
UBC186	*Bitterhof*	134
	Strain No. 164	264, 387
	Strain No. 42	570
	Strain No. 134	438, 630, 649, 966, 990, 1754
RAPD Grass 8	*Strain No. 164*	474
	Bitterhof	788
	Strain No. 134	1705
Pinus 23	*Bitterhof*	310
	Strain No. 134	455
	Japan	
OPA11	*Karafuto No. 1*	538, 1620
	Ezonishiki	305, 593
	Kamishunbetzu	370, 420, 650, 1100
RAPD 377	*Kamishunbetzu*	398
	Shinsei	413, 930
	Ezonishiki	419, 434, 469, 987
	Karafuto No. 1	1744
UBC186	*Karafuto No. 1*	702
	Kamishunbetzu	438
	Shinsei	224, 890,1180, 1404
RAPD Grass 8	*Grignon 48*	181
	Ezonishiki	352
	Karafuto No. 1	732
	Kamishunbetzu	1452
Pinus 23	*Shinsei*	270
	Grignon 48	542, 1200
	Kamishunbetzu	1783

Table 10: Continued

Primers	Accession	Marker size (bp)
	Hungary	
OPA11	*ISZ8*	930
	Keszthelyi Aproszemu Sarga	740
	Balvanska	526, 1044, 1190
	Reatz	2000
RAPD 377	*Balvansk*	357, 1555
	ISZ8	469, 588
UBC186	*Keszthelyi Aproszemu Sarga*	224, 311, 387, 513, 793, 1359
	Balvansk	264, 470, 498, 1121
	Reatz	438, 570
RAPD Grass 8	*Balvansk*	310
	ISZ8	459
	Keszthelyi Aproszemu Sarga	474, 977, 1705
Pinus 23	*Reatz*	270
	Keszthelyi Aproszemu Sarga	1145, 1200, 1520
	Balvansk	1669, 1897
	ISZ8	742
	South Korea	
OPA11	*KAS160_2*	420, 895, 1044, 1480
	KAS133_3	1274
RAPD 377	*KAS131_8*	238, 1208
	KAS133_3	987
	KAS160_2	413
UBC186	*KAS133_3*	206, 311
	KAS131_8	249, 849
	KAS131_9	367
RAPD Grass 8	*KAS131_8*	310, 396
	KAS160_2	474
	PGR 7568	706
	KAS131_9	352
	Netherlands	
OPA11		
RAPD 377	*J_5A*	220,1208
	Ras 20	1096
UBC186	*J_5A*	345
	Ras 20	387, 470, 1118
	No. D.47	793
	No.39	311, 615, 1356
RAPD Grass 8	*No. D.47*	396
	No.39	608
	Ras 20	977
Pinus 23	*No. D.47*	1000, 2000

Table 10: Continued

Primers	Accession	Marker size (bp)
	Russia	
OPA11	*Amurskaja*	265, 483
	Vzlyot	1768
	Bisser	420
	Primorskaja	305, 740
	Ussurijskaja	385, 1100, 1480
RAPD 377	*Ussurijskaja*	443, 516, 686
	Primorskaja	531, 725, 1555, 2000
UBC186	*Bisser*	206, 923
	Primorskaja	311, 470
	Amurskaja	249, 387, 570, 1284
	Ussurijskaja	224
	Vzlyot	345
RAPD Grass 8	*Primorskaja*	459, 474
	Amurskaja	210, 520, 1181, 1452
	Ussurijskaja	1102
Pinus 23	*Ussurijskaja*	270, 1311
	Vzlyot	1323
	Poland	
OPA11	*Bydgoska 052*	260, 1119
	N. 2054	868, 1480, 2000
	Zlotka	370
	Bydgoska 057	804
RAPD 377	*Zlotka*	153
	N. 2054	305
	N. 1954	357, 398, 489, 531
UBC186	*N. 1954*	206, 300, 470
	Bydgoska 052	521, 1180
	Zlotka	367, 570, 849, 1045, 1882
RAPD Grass 8	*N. 1954*	396, 706
	Zlotka	1181
	Bydgoska 052	1394
	N. 2054	1654
Pinus 23	*N. 1954*	455, 1200, 1897
	N. 2054	1145
	Bydgoska 057	1323

Table 10: Continued

Primers	Accession	Marker size (bp)
	Sweden	
OPA11	*744-1*	1854
	698-1-1	1274, 1300
	748-5	740, 1100
RAPD 377	*753-1*	153
	749-2	398, 1284
	744-1	725
UBC186	*698-1-1*	206, 438, 966, 1754
	749-2	286
	744-1	470
	748-5	1121
RAPD Grass 8	*748-5*	310
	744-1	459, 788, 1975
	749-2	948
	698-1-1	706, 1452, 1544
Pinus 23	*753-1*	660
	749-2	775, 1047, 1451
	748-5	1783
	698-1-1	1323

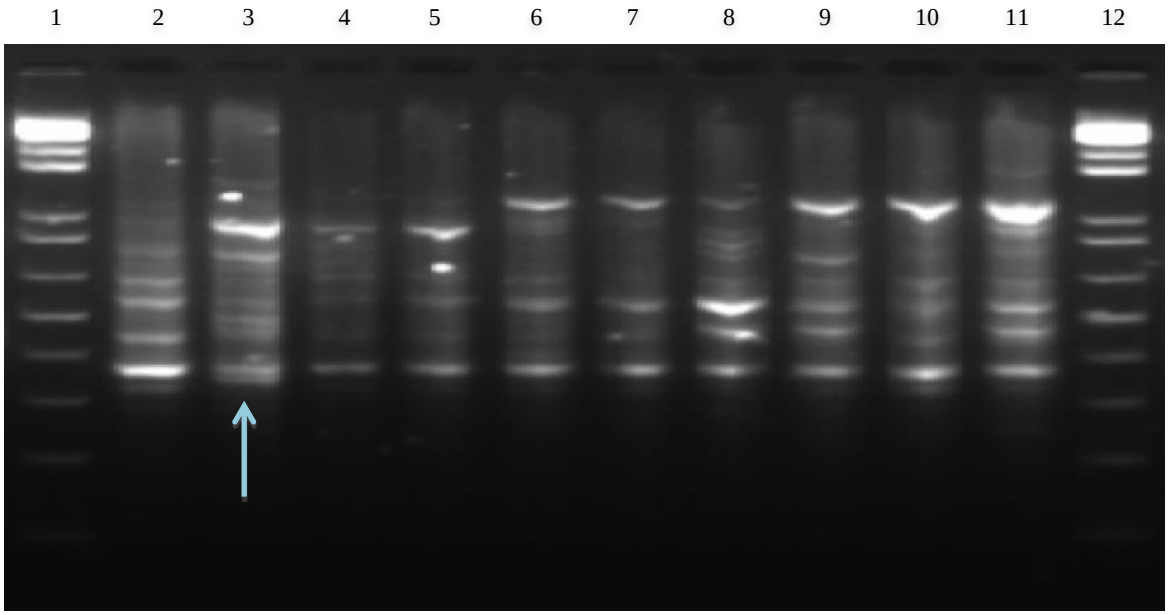


Figure 20: Amplified products using primer ISSR 5. Lane 1 and 12 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 contains accessions from China. The arrow in lane 3 shows a diagnostic marker from *Kao Chien Tao* accession from China

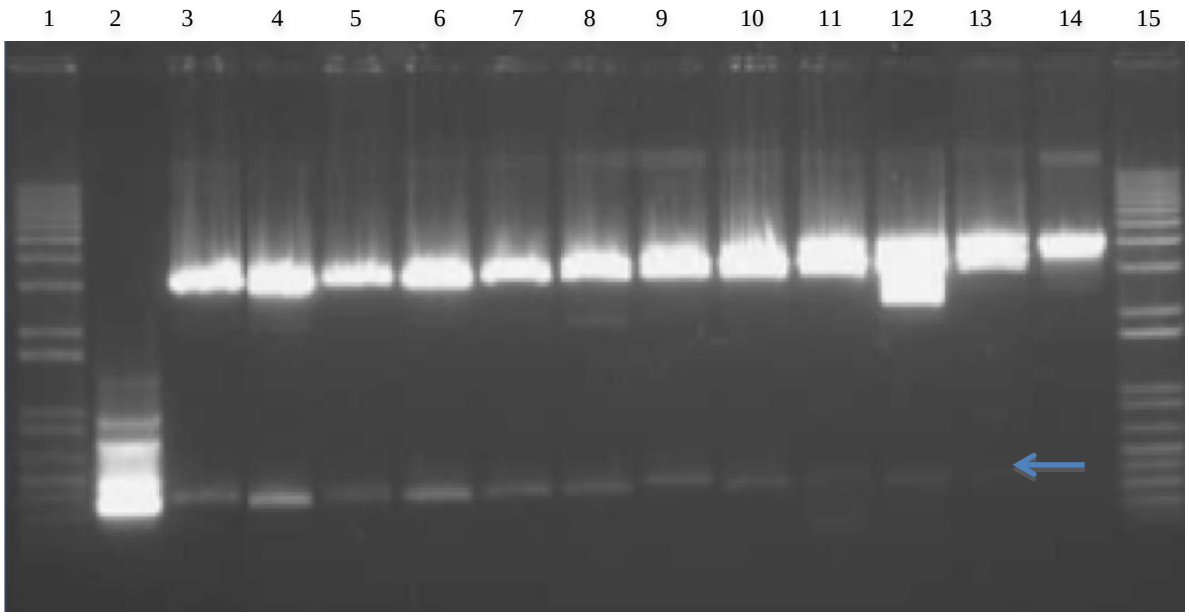


Figure 21: Identification of the diagnostic marker from *Kao Chien Tao* accession from *E.coli* colonies. Lane 1 and 15 contains 1Kb^{+} DNA ladder; lane 2 contains ISSR 5 amplification from accession *Kao Chien Tao* and lane 3 to 14 contains Eco RI digest of plasmid DNA from white colonies

5’

ACGACGACGACGACAACTACGCTCTAGCTGCTTA
ACGCGGCTAGCAGTCGCTAGGGGATGCCTGTAA
ACCCGAAACGACTGTCAGATAGAACAGGATCGC
CGCCAAGTTCGCTGTAGACGTAACGGCTAAAACT
TATACAGCTCGCCCAAAGCACCCTGCCCGTCGG
GCCGCAATGGGTTAACTCAGTAGACAGGGCTAA
GCATGTAGAACCGATAGCGGAGAGCTGGCGGAC
GGGGGTTCAAATCCCCCCGCCTCCACCAAAAAG
AAAGGGCAGCCACTGGCTGCCCTTTTTCTTTGTC
CGTCGTTTAGAACGGGATGTCGTCGTCGTCGT3’

Figure 22: Consensus sequence of a variety specific ISSR fragment 303bp produced by primer ISSR 5.

Table 11: Soybean SCAR 4 Forward and SCAR 4 Reverse Primers

Primer	Sequence (5' to3')	GC content (%)	Primer Length (bp)
SCAR 4 Forward	ACCCGAAACGACTGTCAGAT	50	20
SCAR 4 Reverse	AAGAAAAAGGGCAGCCAGT	45	20

3.4 Discussion

Soybean (*G. max)* breeders evaluate and use the diversity of the germplasm collection to increase knowledge of soybean genetics, to produce soybean for specific food uses, to enhance resistance to disease and to nematodes. Progress will lead to improved soybean germplasm and varieties that survive environmental stress and are competitive in domestic and international markets (White, 2006). However, current breeding practices have led to low level of genetic diversity, which may be ascribed to the emphasis on direct introductions, selection from introduced germplasm and single cross hybrids (some of which shared common parents) in the soybean breeding programs. In addition, *G. max* is a self-pollinating species with no cross-pollination. In fact, the range of genetic diversity within the *G. max* species is not fully characterized, but almost every report on soybean genetic variation has concluded that diversity is lower in comparison with other self -pollinated species. This is attributed to the narrow genetic base of the gene pool available for breeding (Apuya et al., 1988). Therefore, inclusion of more diverse germplasm in the soybean breeding programs may provide the genetic variability necessary to permit continued progress and broad adaptation. A previous report has shown that higher genetic diversity could be found among exotic soybean introductions from different countries (Chowdhury et al., 2002). As discussed in chapter 2, the level of genetic diversity in *G. max* is low. This low level of genetic diversity among *G. max* should facilitate the development of variety-diagnostic or SCAR marker, as a tool to monitor gene transfer among varieties in breeding programs.

In the present study, ISSR and RAPD primers were used to develop markers that distinguish *G. max* varieties within each breeding program. We also identified a 303bp diagnostic marker using ISSR primer 5. This ISSR marker from the Chinese accession *Kao*

Chien Tao from China was absent in all other varieties. Attempts to develop a SCAR marker with designed pair of primers targeting this sequence were not successful. The PCR amplifications did not produce any specific band compared to the control even after optimization. There are several possible reasons for this loss of specificity, such as misinterpretation of ISSR profiles. It is possible that non-homologous sequences of similar size were mistaken for a variety-diagnostic marker (Sanchez de la Hoz et al., 1996; Mehes et al., 2006). Another reason could be the presence of a single nucleotide polymorphism in the annealing region of the ISSR primers (Sanchez de la Hoz et al., 1996; Mehes et al., 2006). Nevertheless, the variety-diagnostic marker is still useful in a breeding program to track the presence of the *Kao Chien Tao* accession in progenies.

One of the difficulties of conversion of random amplified polymorphism and ISSR to a SCAR marker is that desirable cloning of the polymorphic bands cannot be frequently conducted due to the heterogeneous nature of polymorphic bands. In the procedure of converting RAPD and ISSR to SCAR markers, non-targeted sequences can frequently be generated from heterogeneously amplified fragments of similar size with the specific fragment, which might be contained in the polymorphic product determined as one band on the gel images (Lee, 2011). In addition, although the targeted sequence can be obtained, SCAR marker designed from the sequence does not frequently amplify the expected polymorphisms. In most cases, SCAR markers amplify polymorphic products of dominant types. This is because the ISSR technique generally amplifies dominant alleles (Lee, 2011).

There are several reasons why ISSR and RAPD markers should be converted to SCAR in a molecular breeding program. Contrary to ISSR, SCAR primers can amplify only one locus. Thus, the interpretation of the result is far more straightforward as authentication

can be based on the presence and absence of a specific DNA fragment in a given species. Also, the use of longer oligonucleotide primers for SCAR guarantees robust and reliable results. Because of these merits, SCAR has been helpful in differentiating many of the herbal medicines from their close relatives, substituents and adulterants including *Panax* species (Wang et al., 2001), *Astragalus* species (Liu et al., 2008), *Echinacea* species (Adinolfi et al., 2007), *Phyllanthus* species (Theerakalpisut et al., 2008), *Pueraria* species (Devaiah et al., 2008), *Ginger* species (Chavan et al., 2008) *Ipomoea* species (Devaiah et al., 2010) and *Cynanchum* species (Moon et al., 2010).

Success of soybean (*G. max*) breeding is dependent on germplasm availability, genetic variation, selection strategies and resource management. Various crosses between different varieties or germplasm lines are often attempted by breeders to generate increased genetic variation through gene recombination and change of allele frequency in a breeding population in which selection is exercised. Breeding soybean, like other crops, is a numbers game, long term and continuous process, and involves manipulation of genetics of an array of important and complex traits. The strategy to handle a mating scheme and a breeding population structure becomes critical in providing increased potential for genetic superiority, while proper selection and resource management help in improving plant breeding efficiency and success rate (Pengyin et al., 2010).

The premise of this component was to develop variety or accession specific molecular marker for *G. max* breeding programs using ISSR and RAPD markers. Several diagnostic markers were generated for each breeding program. Only one primer revealed a diagnostic marker for *Kao Chien Tao* accession from China, for all accessions from 11 countries. Primers targeting the diagnostic marker sequence were developed. However, the conversion of diagnostic marker to SCAR was not successful. The ISSR diagnostic marker

remains a useful tool for tracking DNA of this accession in hybrid progenies. Further analysis of several other ISSR and RAPD primers is required to achieve the main goal of developing variety-specific markers in the targeted breeding program.

Chapter 4: General Conclusions

Soybean (*Glycine spp*.) is one of the most important crops in the world in terms of total production and usage. It is also among the least diverse species. The main objectives of the present study were 1) to determine differences between ISSR and RAPD marker systems in detecting genetic variation in soybeans and 2) to identify and characterize accession-diagnostic molecular markers in *G. max.*

ISSR and RAPD analyses revealed moderate level of polymorphic loci in soybean accessions from different countries. The highest level of polymorphic loci was observed in accessions from Poland (52%) and the lowest in accessions from the Netherlands (29%). For RAPD marker, the highest level of polymorphism among accessions was 48% for accessions from France and the lowest was 25% in accessions from China. Based on the level of genetic variation detected with both markers (ISSR and RAPD), the soybean gene pools in Netherlands, China, France, and Sweden will benefit the most from introduction of new accessions from other gene banks. In fact, the gene pool of these countries is the least variable genetically.

The genetic distance matrix was generated based on ISSR and RAPD polymorphism data. The ISSR analysis revealed that the accessions from Sweden and Japan were the most genetically closely related with a genetic distance of 0.31. The most genetically distant were the accessions from France and China. Overall, the accessions from different countries were moderately to distantly related. In fact, 82% of genetic distance values were above 0.40 based on ISSR data. The dendrogram revealed that the South Korean accessions formed an out-group. On the other hand, the RAPD analysis showed that accessions from Canada and Netherland were the most genetically closely related with accessions from Russia and South Korea being the most distant. Overall, RAPD data

revealed that the accessions from different countries are closely related with 64% genetic distance values below 0.40. The dendrogram constructed from RAPD data revealed one main cluster. Sweden accessions represented an out-group from the rest with a high degree of confidence.

Attempts were made to develop molecular tools for soybean breeding programs using the 108 accessions characterized in the present study. Several variety–diagnostic markers were identified within each gene pool. One variety-diagnostic marker generated with ISSR 5 primer was identified in the accession *Kao Chien Tao* from China and was absent in DNA from the other 107 accessions.

This marker was sequenced and pair of primers flanking the variety-diagnostic sequence was designed to amplify a SCAR marker. These primers failed to amplify the expected SCAR band. Thus, the diagnostic marker developed was the only tool that could be useful in a breeding program involving the targeted soybean varieties. Several additional ISSR and RAPD primers need to be evaluated to increase the chance of identifing more accession-diagnostic and specific markers.

Reference

Ahmad R., Ferguson L. and Southwick S. (2003). Identification of pistachio (*Pistacia vera* L.) nuts with microsatellite markers. Journal of the American Society for Horticultural Science, 128:pp. 898-903.

Abe J. (2000). The genetic structure of natural populations of wild soybeans revealed by isozymes and RFLPs of mitochondrial DNAs: possible influence of seed dispersal, cross-pollination and demography. In: Oono K., Vaughan D., Tomooka D., Kaga A. and Miyazaki S. The Seventh Ministry of Agriculture, Forestry and Fisheries (MAFF), Japan, International Workshop on Genetic Resources, Ibaraki, Japan. pp. 143-158.

Abe J., Hasegawa A., Fukushi H., Mikami T., Ohara M. and Shimamoto Y (1999) Introgression between wild and cultivated soybeans of Japan revealed by RFLP analysis of chloroplast DNAs. Econ Bot 53:pp.285–291.

Adinolfi B., Chicca A., Martinotti E., Breschi M. and Nieri P. (2007). Sequence Characterized amplified region (SCAR) analysis on DNA from three medicinal Echinacea species. Fitoterap., 78: pp.43-45.

Agarwal M., Shrivastava and Padlh H. (2008). Advances in molecular marker techniques and their application in plant sciences. Plant Cell Reports 27(4):pp.617-31.

Ajibade S., Weeden N. and Chite S. (2000) Inter simple sequence repeat analysis of genetic relationships in the genus Vigna. Euphytica journal. 111:pp.47–55.

Akkaya M., Bhagwat A. and Cregan P. (1992). Length polymorphism of simple sequence repeats DNA in soybean. Genetics 47:pp.1131-1139.

Anderson J. and Bush H. (2011). Soy protein effects on serum lipoproteins: a quality assessment and meta-analysis of randomized, controlled studies. Journal of the American College of Nutrition. 30: pp.79-91.

Apuya N., Frazier B.., Keim P., Jill Roth E. and Lark K. (1988). Restriction length polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merrill. Theoretical and Applied Genetic. 75:pp.889-901.

Armstrong J., Gibbs A., Peakall R. and Weiller G. (1994). The RAPDistance package, version 1.04.

Arslan B. and Okumus A. (2006). Genetic and geographic polymorphism of cultivated tobacco (*Nicotiana tabacum*) in Turkey. Russian Journal of Genetics. 42: pp.667-671.

Bornet B. and Branchard M. (2004). Use of ISSR fingerprints to detect micro satellites and genetic diversity several related Brassica Taxa and Arabidopsis thaliana. Hereditas Journal. 140:pp.245-8.

Bonato A., Calvo E., Geraldi I. and Arias C. (2006). Genetic similarity among soybean (*Glycine max* (L) Merrill) cultivars released in Brazil using AFLP markers. Genetics and Molecular Biology. 29:pp 692-704.

Brown-Guedira G., Thompson J., Nelson R. and Warburton M. (2000). Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. Crop Science. 40:pp.815–823

Chang R. and Chen Y. (1995). Screening for salt tolerance to soybean cultivars of the United States. Soybean Genetic Newsletter. 22:pp.32-42.

Chang R. and Chen Y. (1995). Creative objectives and methods of soybean germplasm in China. Crop Germplasm Resources. 1, pp.2-4.

Chang R., Sun J. and Chen Y. (1997). Collection, conservation and evaluation of soybean genetic resources in China. Chinese Agricultural Science. 1: pp.55-63.

Chang R., Sun J., Qiu L., Chen Y., Li X. and Xu Z. (1999). Collection and conservation of soybean germplasm in ChinaProceedings of Global Soy Forum 99. World Soybean Research Conference VI. pp.172-176.

Chavan P., Waarude D., Joshi K. and Patwardhan B. (2008). Development of SCAR (Sequence-characterized amplified region) markers as a complementary tool for identification of ginger from crude drugs and multicomponent formulations. Biotechnology and Applied Biochemistry. 50: pp.61-69.

Chen Y., Chang R. and Shao G. (1994). Study on the variation of the superoxide dismutase of soybean under salt stress condition. Acta Agrinomica Sinica. 20(3):pp.363-367.

Chen Y. and Nelson R. (1999). Relationship between origin and genetic diversity in Chinese soybean germplasm. Proceedings of World Soybean Research Conference VI, pp. 524.

Chen Y., Wilson R., Fenner G., Kilen T. and Nelson R. (2000). Variation within *Glycine soja* for plant and seed traits. Annual Meeting Abstracts. ASA, CSSA and SSSA. Minneapolis, Minnesota. USA.

Chen Y. (2002) Evaluation of Diversity in *Glycine Soja* and Genetic Relationships within subgenus *Soja.* PhD thesis. Urban, Illinois. pp.156.

Chen Y. (2002). Scholarly articles for Evaluation of Diversity in Glycine Soja and Genetic Relationships Within the Subgenus Soja 22. Soya & Oilseed Bluebook 2001: http://www.soyatech.com/bluebook/statistic/html.

Chowdhury A., Srinives P., Tongpamnak P., Saksoong P. and Chatwachirawong P. (2002). Genetic relationship among exotic soybean introductions in Thailand: Consequence for varietal registration. ScienceAsia 28: pp.227-239.

Congwen S. and Manzhu B. (2006). Genetic Diversity of RAPD Mark for Natural *Davidia involucrata* Populations. pp 95-99.

Costa-Mattioli M., Monpoeho S., Nicand E., Aleman M., Billaudel S. and Ferré V. (2002). Quantification and duration of viraemia during hepatitis A infection as determined by real-time RT-PCR. Journal of *Viral Hepatitis's*. 9: pp.101-106.

Devaiah K. and Venkatasubramaniam P. (2008). Development of SCAR marker for authentication of *Pueraria tuberosa* (Roxb. ex. Willd) Dc. Current science journal. 94 :pp.10.

Devaiah K., Balusubramani S. and Venkatasubramanian P. (2010). Development of randomly amplified polymorphic DNA based SCAR marker for identification of *Ipomoea mauritiana* Jacq (Convolvulaceae). Evidence-Based Complementary and Alternative Medicine. 81 :pp.503-508.

Devos K. and Gale M. (1992). The use of random amplified polymorphic DNA markers in wheat. Theoretical and Applied Genetic. 84:pp.567-572.

Dellaporta S., Wood J. and Hicks J. (1983). A plant DNA minipreparation: version II. Plant Molecular Biology Reporter. 1:pp.19–21.

Diwan N., and Cregan P. (1997). Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theoretical and Applied Genetic. 95: pp.723-733.

Doldi M., Vollmann J. and Lelley T. (1997). Genetic diversity in soybean as determined by RAPD and microsatellite analysis. Plant Breeding. 116: pp.331-335.

Drinic M., Nikolic A. and Peric V. (2008). Cluster Analysis of Soybean Genotypes Based on RAPD Markers. Proceedings of the 43rd Croatian and 3rd International Symposium on Agriculture, Opatija, Croatia, pp.367- 370.

Dwivedi S., Gurtu S., Chandra S., Yuejin W. and Nigam S. (2001). Assesssment of genetic diversity among selected groundnut germplasm. 1. RAPD analysis. Plant Breeding. 120: pp. 345-350.

Ellsworth D., Rittenhouse D., Honeycutt R.(1993). Artifactual variation in random amplified polymorphic DNA banding patterns. Biotechniques . 14:pp. 214-217.

Faheem S., Baloch., Cemal K., Halis A. and Hakan Ö. (2009). Assaying of diversity among soybean (*Glycin max* (L.) Merr.) and peanut (*Arachis hypogaea* L.) genotypes at DNA level. pp.285-303.

Fang Y.,Thomas G., Xu Z., Fang J., Critchley J. and Tomlinson B. (2005). "An affected pedigree member analys is of linkage between the dopamine D2 receptor gene TaqI polymorphism and obesity and hypertension." International Journal of Cardiology 102 (1):pp.111-116.

Fernandez M., Figueiras A. and Benito C. (2002). The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. Theoretical and Applied Genetics, 104: pp.845-851.

Ford B., Ball P. and Ritland K. (1991). Allozyme diversity and genetic relationships among North American members of the short-beaked taxa of *Carex sect. Vesicariae Cyperaceae*. Systematic botany 16(1): pp.116-131.

Fu C. and Chen Y. (1991). Adjust the structure of civil meal and efficient use food legume protein. Grain, oil and food sci-technology. 2:pp.3-6.

Funatsukih H., Ishimoto M., Tsuji H., Kawaguchi K., Hajika M. and Fujno K. (2005). Simple sequence repeat markers linked to a major QTL controlling pod shattering in Soybean. Plant Breeding. 125:pp. 195- 197.

Fukushima D. (2001). Recent progress in research and technology on soybeans. Kikkoman Corp., Noda, Chiba 278-0037.pp. 8-16.

Godwin I., Aitke E. and Smith L. (1997). Application of Inter Simple Sequence Repeat (ISSR) markers to plant genetics. Electrophoresis 18(9): pp.1524-8.

Gizlice Z., Carter T. and Burton J. (1994) Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Science. 34:pp.1143–1151.

Gilbert J., Lewis R., Wilkinson M. and Caligari P. (1999). Developing an appropriate strategy to assess genetic variability in plant germplasm collections. Theoretical and Applied Genetic. 98:pp.1125–1131.

Hiromot D. and Vello N. (1986). The genetic base of Brazilian soybean (*Glycine max* (L.) Merril) cultivars. Brazilian Journal of Genetics, 9: pp.295-306.

Jin Y., He T. and Lu B. (2003). Fine scale genetic structure in a wild soybean (*Glycine soja*) population and the implications for conservation. New Phytologist 159: pp.513–519.

Jin Y., H e T. and Lu B. (2006). Genetic spatial clustering: Significant implications for conservation of wild soybean (*Glycine soja: Fabaceae*). Genetica. 128: pp.41–49.

Juvik G., Bernard R., Chang R. and Cavins J. (1989). Evaluation of the USDA Wild Soybean Germplasm Collection: Maturity Group 000 to IV (PI 65.549 to PI 483.464). U.S. Department of Agriculture. Technical Bulletin. 176: pp. 25.

Karp A., Kresovich S., Bhat K., Ayad W. and Hodgkin T. (1997). Molecular marker: Molecular tools in plant genetic resources conservation: A guide to the technologies new prospects in plant genome analysis. International Plant Genomic Resources Institute. pp. 47.

Kuroda Y., Kaga A., Tomooka N. and Vaughan D. (2006). Population genetic structure of Japanese wild soybean (*Glycine soja*) based on microsatellite variation. Molecular Ecology 15: pp.959–974.

Kuroda Y., Tomooka N., Kaga A., Wanigadeva S. and Vaughan D. (2009). Genetic diversity of wild soybean (*Glycine soja* Sieb. et Zucc.) and Japanese cultivated soybeans [*G. max* (L.) Merr.] based on microsatellite (SSR) analysis and the selection of a core collection. Genetic Resources *and* Crop Evolution. 56:pp.1045-1055.

Tshilenge-Lukanda1 L., Nkongolo K. , Narendrula R., Kalonji-Mbuyi A. and Kizungu R. (2012). Molecular characterization of groundnut (*Arachis hypogaea* L.) accessions from a gene pool: Application of gamma ray radiations. 4: pp. 175-183.

Lee J. and Kaltsikes P. (1973). The application of Mahalanobis's generalized distances to measure genetic divergence in durum wheat. Euphytica journal. 22:pp.124–131.

Li F. (1990). Chinese *G. soja* Collection Catalog. China Agricultural Press, Beijing, China.

Liu T., Lin H. and Wu R. (2008). Identification of Astragalus medicines using SCAR markers. Journal of Food and Drug Analysis. 16:pp.57-62.

Lu H., Li J. S., Liu J. L. and Bernardo R. (2002). Allozyme polymorphisms of maize populations from south western China. Theoretical and Applied Genetics. 104:pp. 119–126.

Mattioni C., Casasoli M., Gonzalez M. and Ipinza R. (2002). Comparison of ISSR and RAPD markers to characterize three Chilean *Nothofagus* variety. Theoretical and Applied Genetics. 104:pp.1064–1070.

Matos M., Pinto-Carnide O. and Benito C. (2001). Phylogenetic relationships among Portuguese rye based on isoenzyme, RAPD and ISSR markers. Hereditas Journal. 134: pp. 299–236.

Mehes M., Nkongolo K. and Michael P. (2007). Genetic variation in *Pinus strobus* and *P. monticola* populations from Canada: development of genome- specific markers. Plant Systematics and Evolution. 267: pp.47-63**.**

Shafie M. and Shafie. (2011). RAPD and ISSR markers for comparative analysis of genetic diversity in wormwood capillary (Artemisia capillaris) from Negeri Sembilan, Malaysia. Journal of Medicinal Plants Research**. 5:** pp. 4426-445.

Moon B., Choo B., Cheon M., Yoon T., Ji Y., Kim B., Lee A. and Kim H. (2010). Rapid molecular authentication of three medicinal plant species, *Cynanchum wilfordii*, *Cynanchum auriculatum*, and *Polygonum multiflorum* (Fallopia multiflorum) by the development of RAPD-derived SCAR markers and multiplex-PCR. Plant Biotechnology Journal. 4: pp. 1-7.

Morgante., Pfeiffer M., Costracurta A. and Oliviero A. (1996). Molecular tools in population and ecological genetics in coniferous trees. Phyton International Journal. 36:pp.129-138.

Mudibu J., Nkongolo K., Mehes -Smith M., and Kalonji -Mbuyi A. (2011). Genetic Analysis of a Soybean Genetic Pool using ISSR Marker: Effect of Gamma Radiation on Genetic Variability. International Journal of Plant Breeding and Genetics. 5: pp.235-245.

Nagaoka T. and Ogihara Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theoretical and Applied Genetics. 94:pp. 597-602.

Narvel J., Fehr W., Chu W., Grant D. and Shoemaker R. (2000). Simple Sequence Repeat Diversity Among Soybean Plant Introductions and Elite Genotypes. Crop Science. 40:pp.1452-1458.

Nei M. (1987). Molecular evolutionary genetics. Columbia University Press, New York, Volume 75, Issue 3, pp.428–429.

Nelson R., Amdor P. and Orf J. (1987). Evaluation of USDA Soybean Germplasm Collection Maturity Group 000 to IV (PI 273.483 to PI 427.107) US Department of Agriculture Technical Bulletin. 1718:pp.267.

Nelson R., Palmer R. and Chen Y. (1993) Genetic diversity in soybean germplasm from central China. The 85th Agronomy Meeting, Cincinnati, OH. USA.

Nkongolo K., Michael P. and Demers T. (2005). Application of ISSR, RAPD, and cytological markers to the certification of *Picea mariana*, *P. glauca*, and *P. engelmannii* trees, and their putative hybrids. Genome 48: pp.302-311.

Prilli R., Mendes-Junior C., Arantes N.E. and Contel E. (2002) Characterization of Brazilian soybean cultivars using microsatellite markers. Genetics and Molecular Biology. 25: pp.185-193.

Qian W., Ge S. and Hong D. (2001). Genetic variation within and among populations of a wild rice *Oryza granulata* from China detected by RAPD and ISSR. Theoretical and Applied Genetics. 102: pp.440-449.

Panthee D., Pantalon V., Saxton A., West D. and Sams C. (2006). Quantitative trait loci for agronomic traits in soybean. Plant breeding journal. 126:pp. 51-57.

Pavlicek A., Hrda S. and Flegr J. (1999). FreeTree - Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Folia Biologica (Praha). 45:pp. 97-99.

Powell W., Morgante M., Andre C., McNicol J., Machray G., Doyle J., Tingey S., Rafalski J. (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Current Biology. 5: pp.1023–1029.

Powell W., Morgante M., Doyle J., McNicol J. and Tingey S. (1996). Genepool variation in genus Glycine. subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics.144:pp.793–803.

Prasad M., Varshney R., Roy J., Balyan H. and Gupta P. (2000). The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theoretical and Applied Genetics.100:pp.584–592.

Queller D., Strassmann J. and Hughes C. (1993). Microsatellites and kinship. Trends in Ecology and Evolution. 8:pp. 285-288.

Raina S., Rani V., Kojima T., Ogihara Y., Singh K. and Devarumath R. (2001). RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome. 44: pp.763-772.

Reddy M., Sarla N. and Siddiq E. (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica journal. 128:pp.9–17.

Rus-Kortekaas W., Smulders M., Arens P. and Vosman B. (1994). Direct comparison of levels of genetic variation in tomato detected by a GACA-containing microsatellite probe and by random amplified polymorphic DNA. Genome. 37:pp. 375-381.

Rongwen J., Akkaya M., Lavi U. and Cregan P. (1995). The use of micro satellite DNA markers for soybean genotype identification. Theoretical and Applied Genetics. 90:pp.43-48.

Roy J., Lakshmikumaran M., Balyan H. and Gupta P. (2004). AFLP-based genetic diversity and its comparison with diversity based on SSR, SAMPL, and phenotypic traits in bread wheat. Biochemical Genetics. 42: pp.43-89.

Sambrook J., Fritsch E. and Maniatis T. (1989). Molecular Cloning: A Laboratory Manual, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sangiri C., Kaga A., Tomooka N., Vaughan D. and Srinives . (2007). Genetic diversity of the mungbean (*Vigna radiata*, Leguminosae) genepool on the basis of microsatellite analysis. Australian Journal of Botany. 55:pp.837-847.

Sanchez de la Hoz M., Da J., Loarce Y. and Ferrer E. (1996). Simple sequence repeat primers used in polymerase chain reaction amplifications to study genetic diversity in barley. Genome. 39:pp.112-11.

Semagn K., Bjornstad and Ndjiondjop M. (2006). An overview of molecular markers methods for plants. African Journal of Biotechnology 5: pp.2540-2568.

Shao G., Chang R. and Chen Y.(1994). Study on inheritance of salt tolerance in soybean. Proceedings of World Soybean Research Conference V, Chiang Mai, Thailand. pp.581.

Shao G., Chang R. and Chen Y.(1993). Study advance of salt tolerance on soybean. Soybean Sciences. 12: pp. 211-214.

Sharma A., Namdeo A. and Mahadik K. (2008). Molecular markers: New prospects in plant genome analysis. Pharmacognosy Reviews 2: pp. 23-34.

Shoemaker R., Cregan P. and Vodkin L. (2004). Soybean Genomics. In: Soybeans: Improvement, Production and Uses, Boerma, H.R. & Specht, J.E. (Eds.), Wisconsin, USA. pp.235-263.

Souframanien J. and Gopalakrishna T. (2004). Comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers. Theoretical and Applied Genetics. 109: pp.1687-1693.

Sudaric A., Vrataric M., Duvnjak T., Sudar R. and Rajcan I. (2005). Application of molecular markers in soybean breeding program at the Agricultural Institute Osijek. Sjemenarstvo. 22:pp.246-247.

Theerakulpisut P., Kanawpee N., Maensiri D., Bunnag S.and Chantaranothai P. (2008). Development of species-specific SCAR markers for identification of three medicinal species of Phyllanthus. Journal of Systematics and Evolution. 46: pp. 614-21.

Ude G., Kenworthy W., Costa J., Cregan P. and Alvernaz J. (2003). Genetic diversity of soybean cultivars from China, Japan, North America and North American Ancestral lines determined by amplified fragments length polymorphism. Crop Science. 43: pp.1858-1867.

Wang J., Ha W., Ngan F., But H. and Shaw P. (2001). Application of sequence characterized amplified region (SCAR) analysis to authenticate Panax species and their adulterants. Planta Medica. 67: pp.781-783.

Wendy W. (2006). Plant Breeding, Genetics, and Cytogenetics. http://cropsci.illinois.edu/content/plant-breeding-genetics-and-cytogenetics.

Whitehouse W. (1957). The pistachio nut. A new crop for the Western United States. Economical Botany. 11: pp.281-321.
http://www.mayoclinic.com/health/soy/NS_patient-soy.

Williams J., Kubelik A., Livak K., Rafalski J. and Tingey S. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 18: pp.6531-6535.

Xu D., Abe J., Gai J. and Shimamoto Y (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated cultivated soybeans. Theoretical and Applied Genetics. 105:pp.645–653.

Ye Y., Wang Z. and Zhuo S. (2012). Soy germ isoflavones improve menopausal symptoms but have no effect on blood lipids in early postmenopausal Chinese women: a randomized placebo-controlled trial. Menopause. 19: pp.791-8.

Yeh F. and Boyle T. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian Journal of Botany. 129: pp. 157.

Zietkiewicz E. Rafalski and Labuda D. (1994). Genome fingerprinting by simple sequence repeat (SSR)- anchored polymerase chain reaction amplification. Genomics. 20: pp.176-183.

Zhang Z. and Chen Y. (1988). Main agronomic and economic characters of peas in China. Crop Germplasm Resources. 2: pp. 6-9.

Zhang X., You G., Dong Y., Jia J., Liu X., Shang X., Liu S. and Cao Y. (2006). Genetic diversity in Chinese modern wheat varieties revealed by microsatellite markers. Science China Life Sciences 49: pp.218–226.

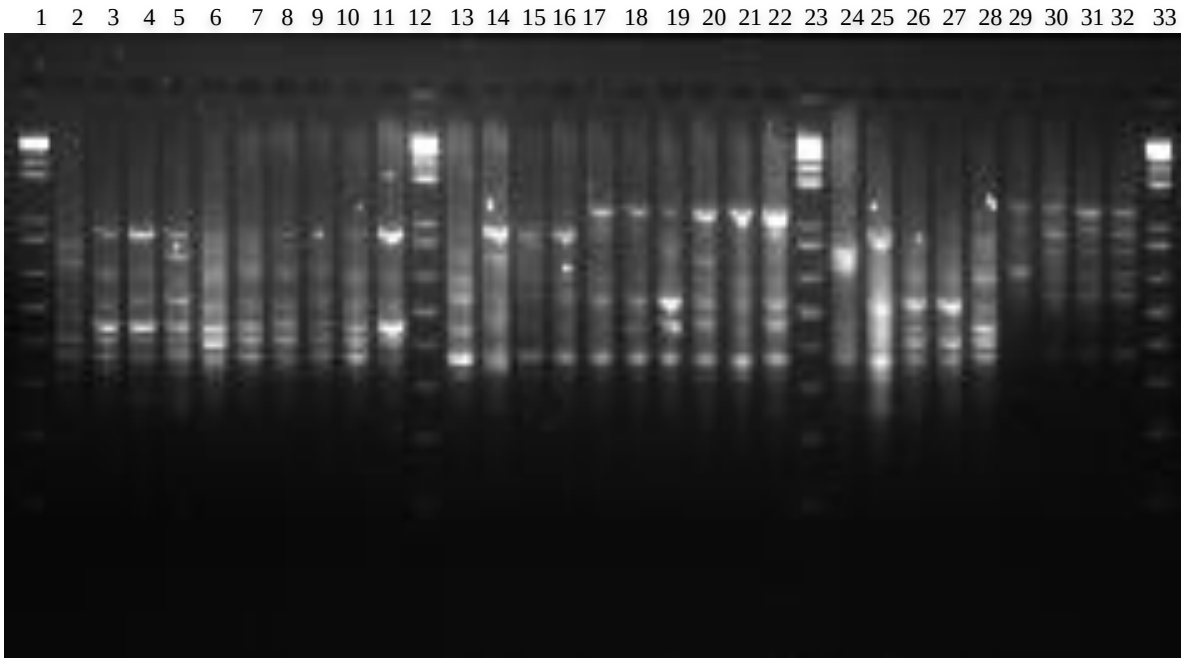
Appendices

Appendix 1: List of ISSR primers

ISSR Primers				
ISSR 829	UBC 837	ISSR HB 13	UBC 844	ISSR 829
SC ISSR 6	UBC 873	ISSR8	UBC827	SC ISSR 6
UBC 835	UBC 849	ISSR Echt 6	ISSR17878	UBC 835
ISSR Echt4	UBC 825	ISSR 4	UBC 879	ISSR Echt4
ISSR 7	ISSR Echt5	UBC 818	ISSR10	ISSR 7
ISSR 5	ISSR 3	UBC 812	ISSR17899	ISSR 5
ISSR 49	ISSR 1	ISSR9	ISSR Echt 7	ISSR 49
ISSR HB 15	UBC 823	ISSR 112	UBC804	ISSR HB 15

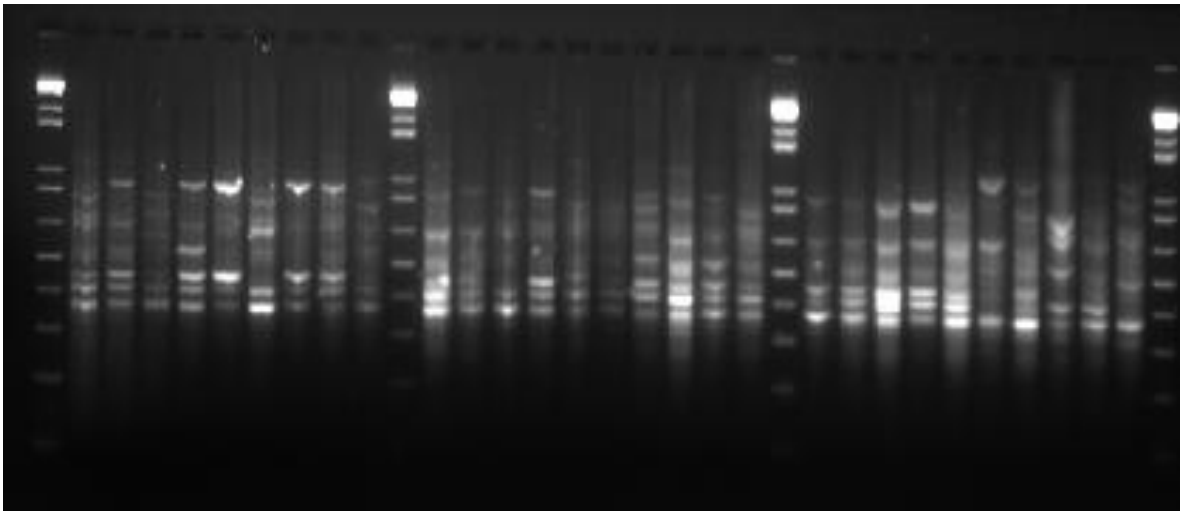
Appendix 2: List of RAPD primers

RAPD Primers			
Grasse 8	UBC186	RAPD 49	OPA 11
Grasse 2	OPA4	RAPD 201	UBC 337
OPA46	OPA 26	OPA2	PINUS 23

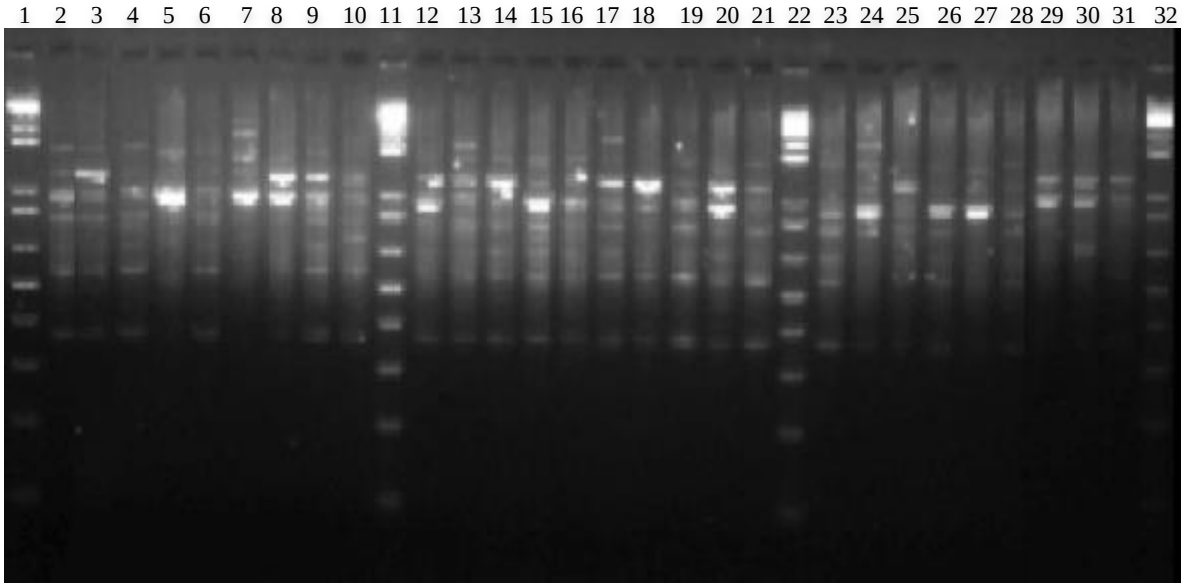


Appendix 3: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 5. Lanes 1, 12, 23, and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Germany; lanes 13 to 22 contain accessions from Japan and lanes 24 to 32 contain accessions from Hungary.

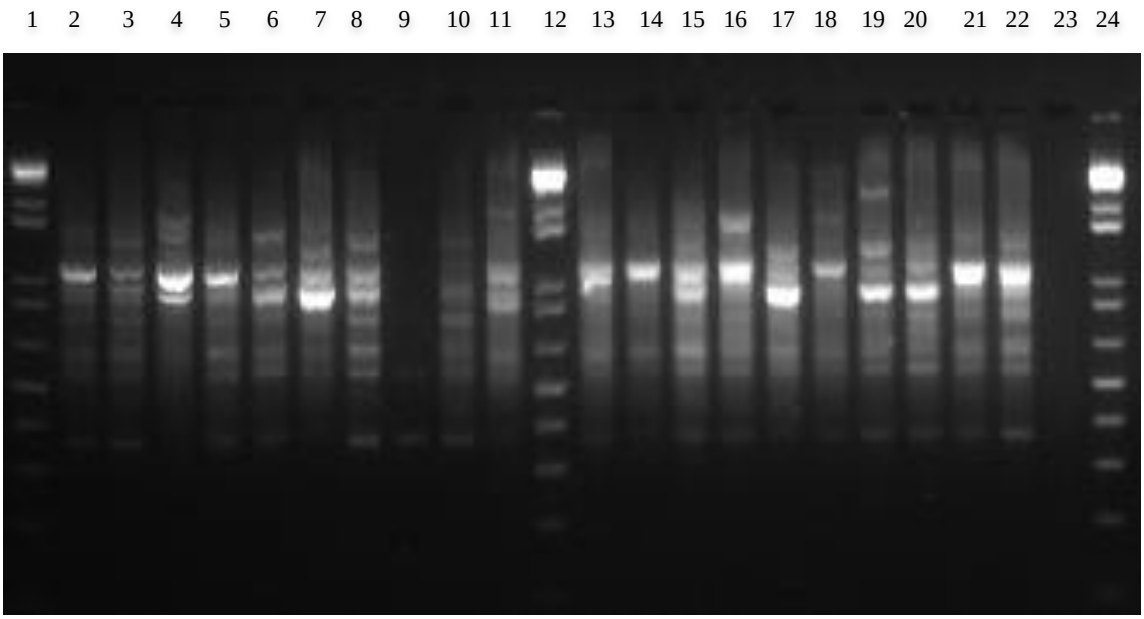
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33



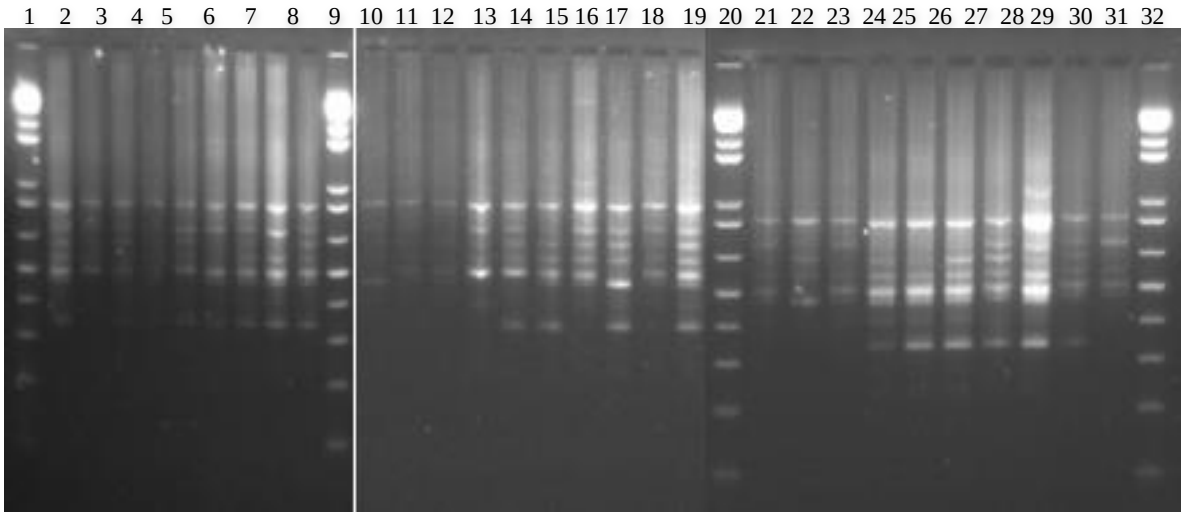
Appendix 4: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 5. Lanes 1, 11, 22 and 33 contain 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Canada; lanes 12 to 21 contain accessions from China and lanes 23 to 32 contain accessions from France.



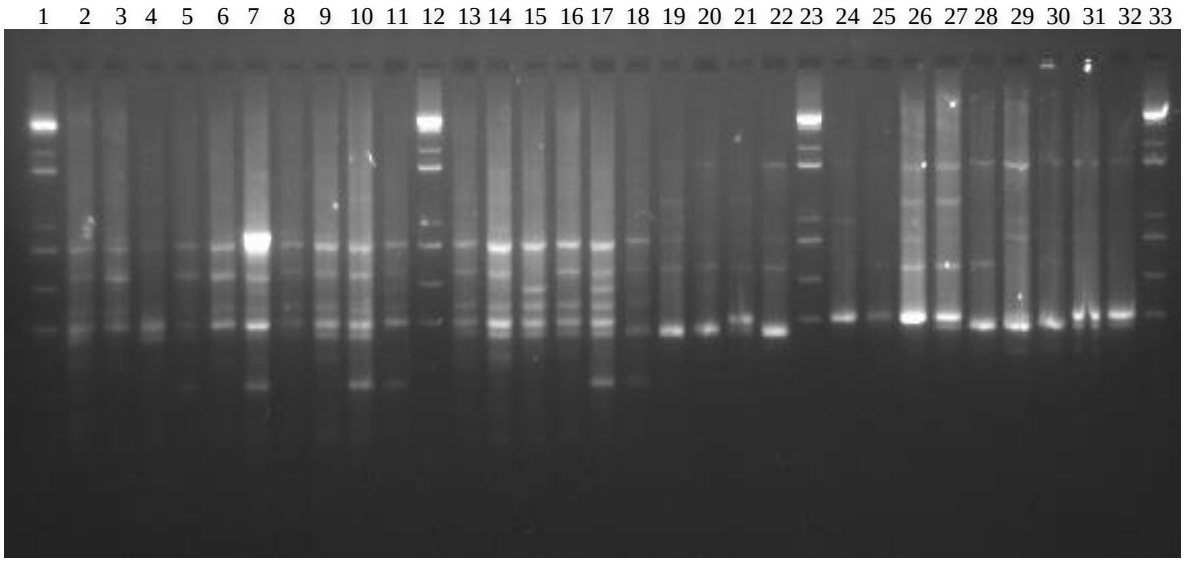
Appendix 5: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 5. Lanes 1, 11, 22, and 32 contain 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from South Korea; lanes 12 to 21 contain accessions from Netherlands and lanes 23 to 31 contain accessions from Russia.



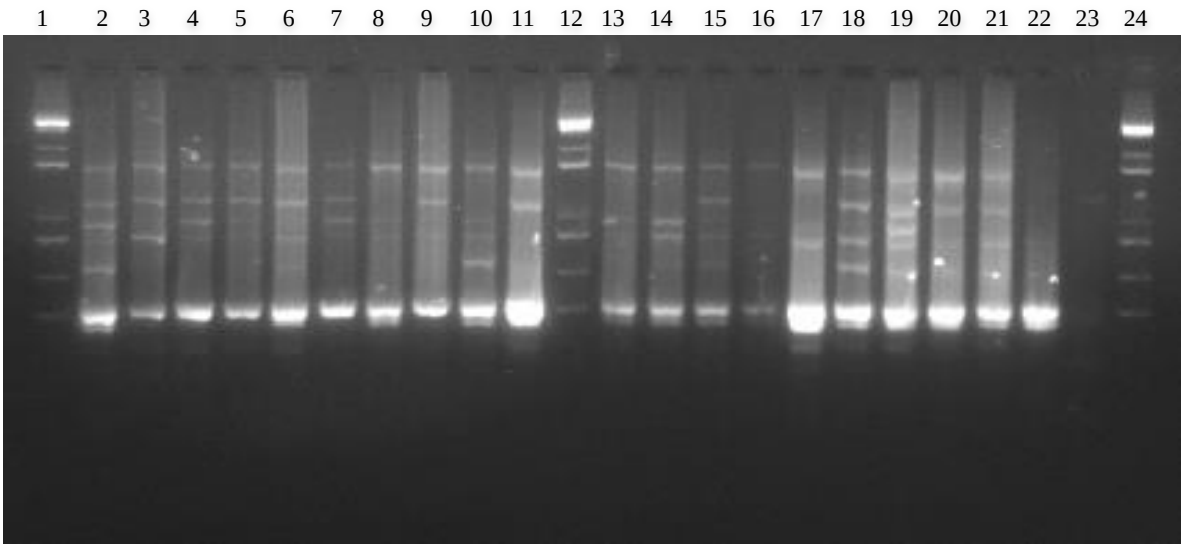
Appendix 6: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 5. Lanes 1, 12, and 24 contain 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Poland and lanes 12 to 21 contain accessions from Sweden.



Appendix 7: ISSR amplification of Soybean (*Glycine max*) accessions with primer ISSR 849. Lanes 1, 11, 22, and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 10 represents accessions from Canada; lanes 12 to 21 contain accessions from China and lanes 23 to 32 contain accessions from France.

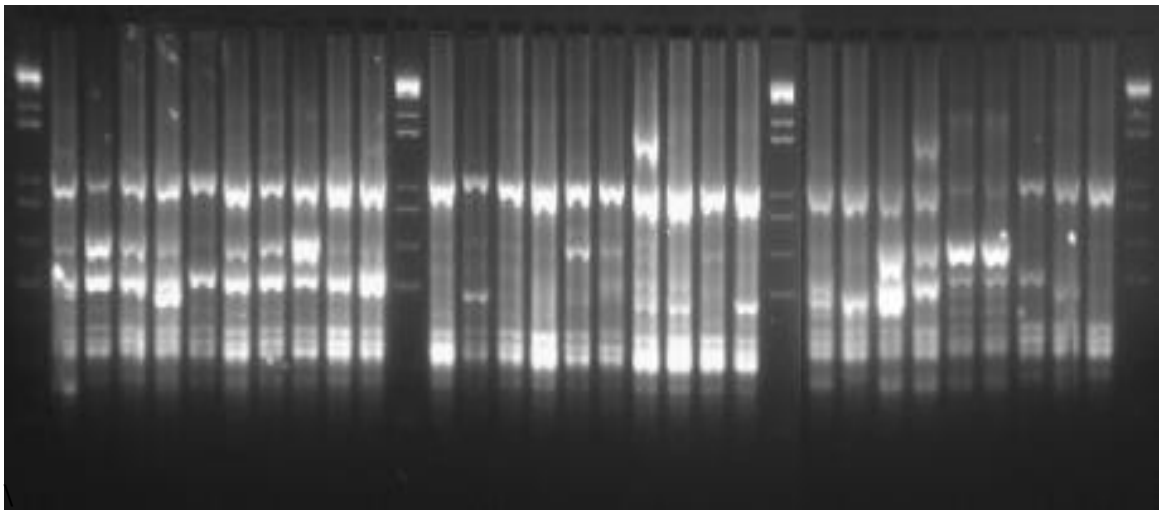


Appendix 8: ISSR amplification ISSR amplification of Soybean (*Glycine max*) accessions with primer ISSR 849. Lanes 1, 12, 23 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from South Korea; lanes 13 to 22 contain accessions from Netherlands and lanes 23 to 32 contain accessions from Russia.

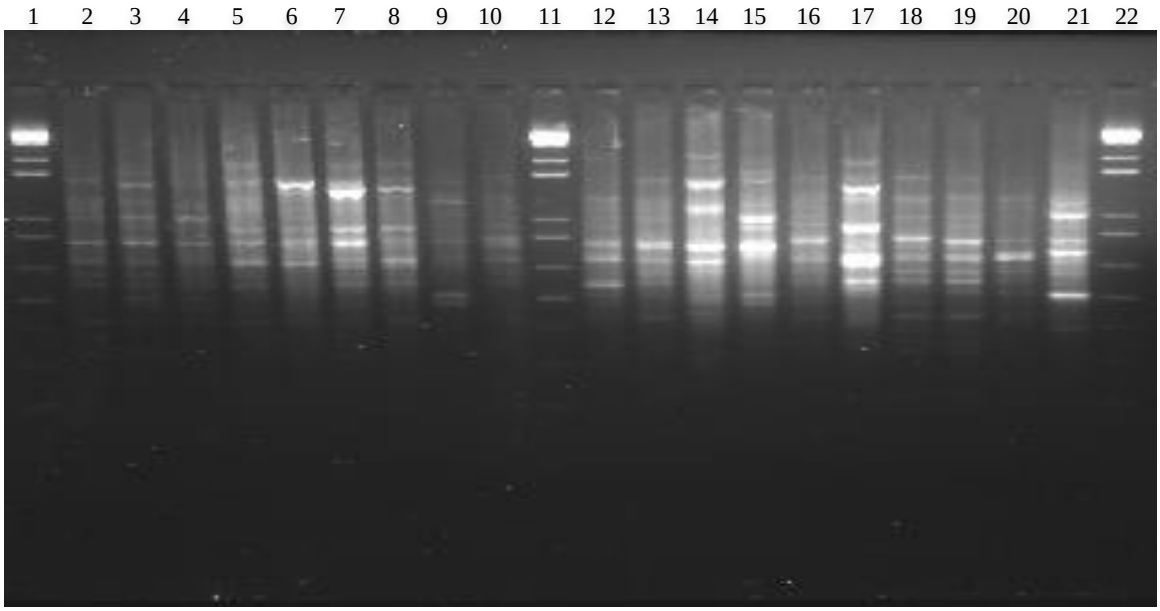


Appendix 9: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 849. Lanes 1, 12 and 24 contain 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Poland and lanes 12 to 21 contain accessions from Sweden.

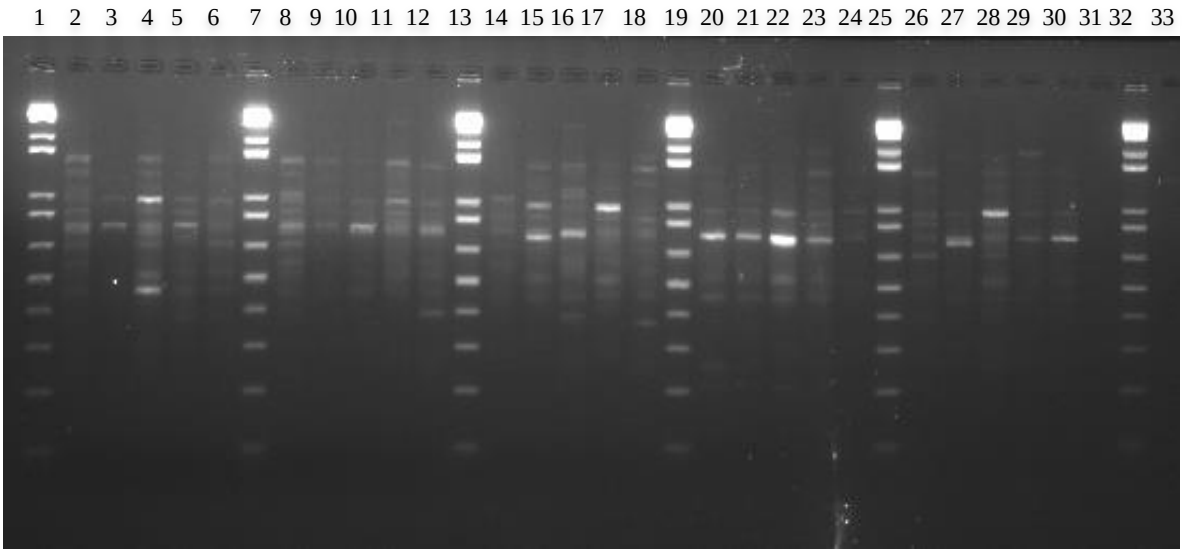
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33



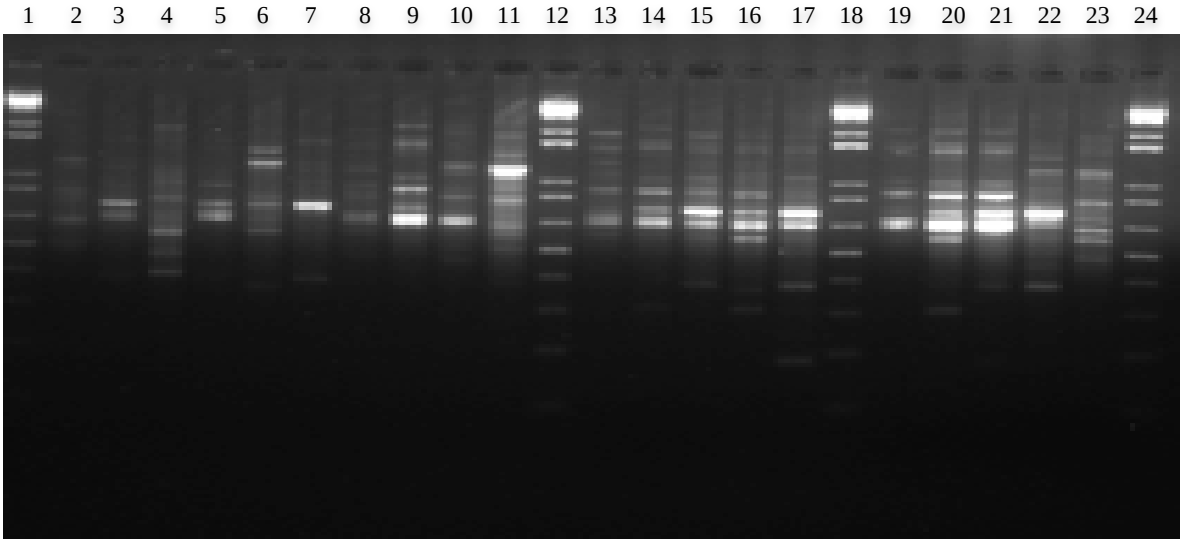
Appendix 10: ISSR amplification of Soybean (*Glycine max)* accessions with primer ISSR 873. Lanes 1, 12, 23 and 33 contains 1Kb^{+} DNA ladder. Lanes 2 to 11 represents accessions from Germany; lanes 13 to 22 contain accessions from Japan and lanes 24 to 32 contain accessions from Hungary.



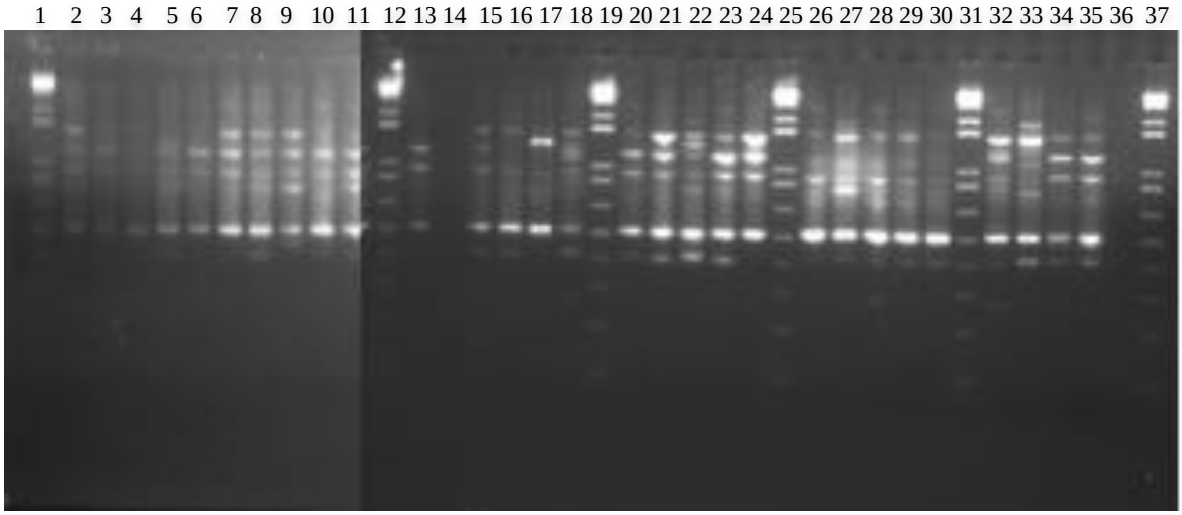
Appendix 11: ISSR amplification of Soybean (*Glycine max*) accessions with primer ISSR Echt 6. Lanes 1, 11 and 22 contain 1Kb^{+} DNA ladder. Lanes 2 to10 represents accessions from Russia and lanes 12 to 21 contain accessions from Poland.



Appendix 12: RAPD amplification of Soybean (*Glycine max*) accessions with primer OPA 11. Lanes 1, 7, 1 3, 19, 25 and 32 contain 1Kb^{+} DNA ladder. Lanes 2 to 6 represents accessions from South Korea; lanes 8 to 12 contain accessions from Netherlands; Lanes 14 to18 contain accessions from Russia; lanes 20 to 24 contain accessions from Poland and lanes 26 to 30 contain accessions from Sweden



Appendix 13: RAPD amplification of Soybean (*Glycine max*) accessions with primer Grasse 8. Lanes 1, 12, 18 and 24 contain 1Kb^{+} DNA ladder. Lanes 2 to 6 represents accessions from Canada; lanes 7 to 11 contain accessions from China; lanes 13 to 17 contain accessions from Germany; lanes 19 to 23 contain accessions from Japan.



Appendix 14: RAPD amplification of Soybean (*Glycine max*) accessions with primer Pinus 23. Lanes 1, 12, 19, 25, 31 and 37 contain 1Kb^{+} DNA ladder. Lanes 2 to 11 represent lanes from Canada; lanes 7 to 11 contain accessions from China; lanes 13 to 18 contain accessions from France; lanes 20 to 24 contain accessions from Germany; lanes 26 to 30 contain accessions from Japan and lanes 32 to 36 contain accessions from Hungary